Reformulating the Dual Graphs of CSPs to Improve the Performance of Relational Neighborhood Inverse Consistency

Robert J. Woodward[1] **Shant Karakashian**[1] **Berthe Y. Choueiry**[1] **Christian Bessiere**[2]

[1]Constraint Systems Laboratory, University of Nebraska-Lincoln, USA
{rwoodwar|shantk|choueiry}@cse.unl.edu

[2]LIRMM-CNRS, University of Montpellier, France
bessiere@lirmm.fr

Abstract

Freuder and Elfe (1996) introduced Neighborhood Inverse Consistency (NIC) as a new local consistency property for binary Constraint Satisfaction Problems (CSPs). Two advantages of the algorithm for enforcing NIC is that it automatically adapts its filtering power to the local connectivity of the network and has insignificant space overhead. However, studies on binary CSPs have shown that enforcing NIC is not effective on sparse graphs and too costly on dense graphs. In (Woodward et al. 2011), we introduced an algorithm for enforcing Relational Neighborhood Inverse Consistency (RNIC), which is an extension of NIC to non-binary CSPs. In this paper, we discuss how we enhance the propagation effectiveness of our algorithm and reduce its computational cost by reformulating the dual graph of the CSP. For that purpose, we describe two reformulation techniques that modify the topology of the dual graph without affecting the solution set of the problem. We present the two reformulations and their combinations, and discuss their effects on the consistency property enforced by the algorithm. We also describe a selection policy that nicely ties together the various components of our approach in a consistent, adaptive framework. Finally, we show that our automated selection policy outperforms all approaches in a statistically significant manner.

1 Introduction

An important result in Constraint Processing (CP) ties the tractability[1] of a Constraint Satisfaction Problem to the level of consistency that it satisfies. Solving difficult problems often requires enforcing higher order consistency, which typically requires the use of more costly algorithms in time and/or in space. Freuder and Elfe (1996) introduced Neighborhood Inverse Consistency (NIC) for binary CSPs as a particularly promising consistency property because: (1) Enforcing it is light in terms of space requirements (inverse consistency is enforced by filtering the variables domains); and (2) It focuses the attention on where a variable's value most tightly interacts with the problem, namely its neighborhood. Despite its promise and filtering effectiveness, NIC remains relatively unexploited because the algorithm for enforcing it is too costly in terms of processing time, which prevented its use on dense networks or in a lookahead scheme during backtrack search.

In (Woodward et al. 2011), we generalized NIC to Relational Neighborhood Consistency (RNIC) for non-binary CSPs. Although, Bacchus et al. (2002) had already identified the same property as RNIC to hold when the arc-consistency property holds in on the dual graph[2] of the CSP, they do not provide a practical algorithm for enforcing it, study its usefulness in practice, or compare to any consistency properties other than arc consistency, all of which we examine in this paper.

This paper is structured as follows. Section 2 reviews background information about CSPs. Section 3 introduces RNIC and an algorithm for enforcing it on the dual encoding of the CSP. Section 4 discusses three variations of RNIC obtained by removing redundant edges in the dual graph and/or triangulating the considered graph, and a strategy for deciding which of the four properties to enforce. The goal of this deliberation is to reduce computational cost and/or strengthen propagation depending on the topology of the dual graph. Section 5 reviews the state of the art in relational consistency. Section 6 discusses our experimental results, we compare the performance of the resulting mechanisms, on difficult benchmark problems, with that of GAC2001 (Bessière et al. 2005) and the recently introduced algorithms for m-wise consistency (i.e., wR($*$,m)C for $m = 2$, 3, 4 of (Karakashian et al. 2010)). Finally, Section 7 discusses the extension of our approach to relations specified as conflicts or in intension and concludes this paper with directions for future research.

2 Background

A Constraint Satisfaction Problem (CSP) is defined by $\mathcal{P} = (\mathcal{V}, \mathcal{D}, \mathcal{C})$ where $\mathcal{V}$ is a set of variables, $\mathcal{D}$ is a set of domains, and $\mathcal{C}$ is a set of constraints. Each variable $V_i \in \mathcal{V}$ has a finite domain $D_i \in \mathcal{D}$, and is constrained by a subset of the constraints in $\mathcal{C}$. Each constraint $C_i \in \mathcal{C}$ is specified by a relation R_i defined on a subset of the variables, called the scope of the relation and denoted $scope(R_i)$. Given a relation R_i, a tuple $\tau_i \in R_i$ is a vector of allowed values for the variables in the scope of R_i. Solving a CSP corresponds to finding an assignment of a value to each variable such that all the constraints are satisfied.

[1]The tractability of a problem is the ability to solve it in time polynomial in the size of the input, which, in the case of the CSP, is the number of variables.

[2]The dual graph is defined in Section 2.

2.1 Graphical representations

A binary CSP is represented by its *constraint graph* where the vertices are the variables of the CSP and the edges represent the constraints. A non-binary CSP is similarly represented by its *hypergraph* where the hyperedges represent the non-binary constraints. Another graphical representation of a non-binary CSP is the *primal graph* where the vertices are the CSP variables and edges connect every two vertices corresponding to variables in the scope of a relation (Dechter 2003). $\mathrm{Neigh}(V_i)$ denotes the set of variables that are adjacent to V_i in the constraint graph of a binary CSP and the primal graph of a non-binary CSP. The dual encoding of a CSP $\mathcal{P}$ is a binary CSP whose variables are the relations of $\mathcal{P}$, their domains are the tuples of those relations, and the constraints enforce *equalities* over the shared variables. The representation as a graph of this encoding is the *dual graph* of the CSP. $\mathrm{Neigh}(R_i)$ denotes the set of relations adjacent to a relation R_i in the dual graph. Figure1, illustrates the hyper, primal, and dual graphs of a small non-binary CSP where $\mathcal{V} = \{A, \ldots, F\}$ and the relations are $R_1, \ldots, R_6$.

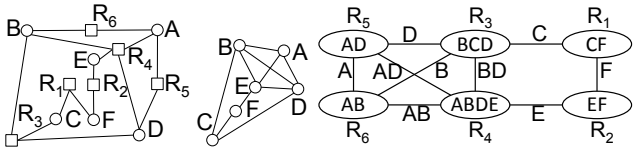

Figure 1: Hyper, primal, and dual graphs of a small CSP.

2.2 Consistency properties & algorithms

CSPs are in general $\mathcal{NP}$-complete and solved by search. To reduce the severity of the combinatorial explosion, they are usually 'filtered' by enforcing a given local consistency property (Bessiere 2006). One common such property is Generalized Arc Consistency (GAC). A CSP is GAC iff, for every relation, any value in the domain of any variable in the scope of the relation can be extended to a tuple satisfying the relation. Our work extends, to non-binary CSPs, the local consistency property known as Neighborhood Inverse Consistency (NIC) introduced, for binary CSPs, in (Freuder and Elfe 1996). NIC ensures that every value in the domain of a variable can be extended to a solution of the subproblem induced by the variable and the variables in its neighborhood. Algorithms for enforcing GAC and NIC typically operate by filtering the domains of the variables. Pairwise consistency[3] (Janssen et al. 1989), R($*$,m)C (Karakashian et al. 2010) and RNIC are consistency properties of the dual graph of the CSP. The algorithms for enforcing them typically operate by filtering the constraint definitions.

In order to compare the various consistency properties discussed in this paper we use the terminology introduced in (Debruyne and Bessière 1997). Given two consistency properties p and p',

- p is *stronger* than p' if, in any CSP where p holds, p' also holds.
- p is *strictly stronger* than p' if p is stronger than p' and there exists at least one CSP in which p' holds but p does not.
- p and p' are *equivalent* when p is stronger than p' and vice versa.
- Finally, p and p' are *incomparable* when there exists at least one CSP in which p holds but p' does not, and vice versa.

In practice, when a consistency property is stronger (respectively, weaker) than another, enforcing the former never yields less (respectively, more) pruning than enforcing the latter on the same problem.

3 Relational NIC

The algorithm for enforcing NIC on binary CSPs of (Freuder and Elfe 1996) was tested in a preprocessing step to backtrack search on instances whose constraint density[4] did not exceed 4.25%. Despite its pruning power and light space overhead, NIC received relatively little attention in the literature, likely because of the prohibitive cost of the algorithm for enforcing it. Below, we introduce RNIC, a generalization of NIC to non-binary CSPs, and characterize it. Then, we describe and analyze an algorithm for enforcing it.

3.1 Defining RNIC

Definition 1 *A relation R_i is said to be* RNIC *iff every tuple in R_i can be extended to the variables in $\bigcup_{R_j \in \mathrm{Neigh}(R_i)} scope(R_j) \setminus scope(R_i)$ in an assignment that simultaneously satisfies all the relations in* $\mathrm{Neigh}(R_i)$. *A network is* RNIC *iff every relation is* RNIC.

Informally, every tuple τ_i in every relation R_i can be extended to a tuple τ_j in each $R_j \in \mathrm{Neigh}(R)$ such that together all those tuples are consistent with all the relations in $\mathrm{Neigh}(R_i)$. Like $R(*, m)C$, RNIC can be enforced by filtering the existing relations and without introducing any new relations to the CSP. A straightforward algorithm for enforcing RNIC applies the following operation to every relation R_i in the problem until quiescence:

$$R_i \leftarrow \pi_{scope(R_i)}(\bowtie_{R_j \in \{R_i\} \cup \mathrm{Neigh}(R_i)} R_j) \qquad (1)$$

where π and $\bowtie$ are the relational operators project and join, respectively. The space requirement of this algorithm is prohibitive in practice because it requires storing the join of $R_i \cup \mathrm{Neigh}(R_i)$, which is not necessary as we argue in Section 3.4. For the example of Figure 1, RNIC examines the six subproblems induced on the dual graph by each relation and its neighborhood as listed below:

1. For R_1, $\mathrm{Neigh}(R_1) = \{R_2, R_3\}$.
2. For R_2, $\mathrm{Neigh}(R_2) = \{R_1, R_4\}$.
3. For R_3, $\mathrm{Neigh}(R_3) = \{R_1, R_4, R_5, R_6\}$.
4. For R_4, $\mathrm{Neigh}(R_4) = \{R_2, R_3, R_5, R_6\}$.
5. For R_5, $\mathrm{Neigh}(R_5) = \{R_3, R_4, R_6\}$.
6. For R_6, $\mathrm{Neigh}(R_6) = \{R_3, R_4, R_5\}$.

Generally speaking, the number of induced subproblems to be considered is equal to e, where e is the number of relations in the CSP; and the size of the largest subproblem is equal to $\delta + 1$, where δ is the degree of the dual graph.

[3] Pairwise consistency (Janssen et al. 1989) is equivalent to R($*$,2)C (Karakashian et al. 2010).

[4] The constraint density of a binary CSP is equal to $\frac{2e}{n(n-1)}$, where e is the number of constraints and n the number of variables.

3.2 Comparing RNIC and R(∗,m)C

In (Karakashian et al. 2010), we introduced the property R(∗,m)C with $m \geq 2$, which ensures that every tuple in every relation can be extended in a consistent assignment to every combination of $m-1$ relations in the problem. For the example shown in Figure 1, R(∗,2)C must verified on 9 combinations of two relations. Generally speaking the number of induced subproblems to be considered is $\mathcal{O}(e^m)$; and the size of the largest subproblem is equal to m. We compare RNIC with R(∗,m)C, which is defined for $m \geq 2$.

1. RNIC is strictly stronger than R(∗,m)C, $m \leq 3$ (see Theorem 2 in (Woodward et al. 2011)).
2. R(∗,m)C with $m \geq \delta + 1$, where δ is the degree of the dual graph, is strictly stronger than RNIC (see Theorem 3 in (Woodward et al. 2011)).
3. For $4 \leq m \leq \delta$, R(∗,m)C and RNIC are not comparable (see Theorem 4 in (Woodward et al. 2011)).

Figure 2 illustrates the above first three assertions. Two interesting structures of the dual graphs, trees and cycles, are such that several relational consistency properties collapse to R(∗,2)C, which is the weakest of them all:

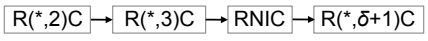

Figure 2: Comparing RNIC with R(∗,m)C.

Theorem 1 *RNIC, R(∗,2)C, and R(∗,m)C are equivalent on any dual graph that is tree structured or is a cycle of length* $\geq maximum(4, m+1)$.

Proof: By straightforward generalization of Theorem 5 in (Woodward et al. 2011). □

The theorem applies for a tree of any degree. As for the cycle, it must be length at least $m+1$ for $m \geq 3$. This last theorem is important because it identifies structural configurations where the relational consistency properties RNIC and R(∗,m)C collapse to their weakest version, that is R(∗,2)C. In Section 4 we propose reformulating the dual graph of the CSP to allow RNIC to overcome this obstacle.

3.3 Comparing RNIC and domain filtering

In practice, after enforcing RNIC on a CSP (by filtering the relations), the domains of the variables are updated accordingly in order to reduce the search effort. It is important to note that variable domains can be updated by simply projecting the filtered relations on the variables. Interestingly, these domain reductions do not break the RNIC property.

Theorem 2 *If a network is RNIC, domain filtering by GAC cannot enable further constraint filtering by RNIC.*

Proof: Similar to proof of Theorem 1 in (Karakashian et al. 2010). □

Following the terminology of (Bessière, Stergiou, and Walsh 2008), the property of a CSP where RNIC holds and where the domains agree with the constraints is denoted RNIC+GAC. Although formally correct, we find this notation confusing because it may incorrectly suggest the need to enforce GAC, which is in general more expensive than (simply and without looping) projecting the relations on the domains. For that reason, we choose to denote this property instead RNIC+DF (i.e., RNIC followed by domain filtering).

Theorem 3 *NIC (on a binary CSP) and RNIC+DF (on the dual graph of the same binary CSP) are not comparable.*

Proof: In Figure 3, the CSP is NIC but not RNIC+DF. RNIC removes the tuples in $\{(0,2),(2,2)\}$ from R_0,

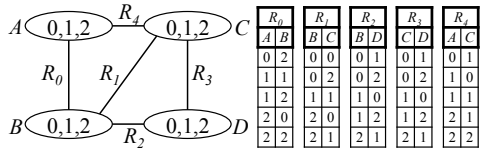

Figure 3: The binary CSP is NIC but not RNIC+DF.

$\{(0,0),(1,2)\}$ from R_1, $\{(0,2)\}$ from R_2, $\{(0,2)\}$ from R_3, and $\{(0,1),(2,1)\}$ from R_4. Therefore, RNIC+DF removes the value 0 from A. In Figure 4, the CSP is RNIC+DF but not NIC. NIC removes the value 0 from D. □

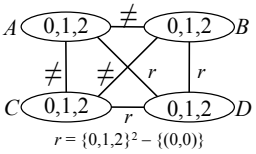

Figure 4: Binary CSP is RNIC+DF but not NIC.

The *singleton* variant of a given consistency property guarantees that the assignment of every value in the domain of a variable yields a CSP where the consistency property holds (Debruyne and Bessière 2001). Singleton consistencies have been studied mainly for arc consistency (SAC) and generalized arc consistency (SGAC).

Theorem 4 *SGAC on a non-binary CSP and RNIC+DF on the corresponding dual graph are not comparable.*

Proof: In Figure 5, the CSP is RNIC+DF but not SGAC. SGAC empties all variables domains. In Figure 6, taken

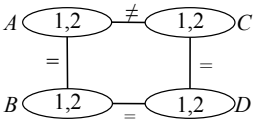

Figure 5: The CSP is RNIC+DF but not SGAC.

from (Debruyne and Bessière 2001), the CSP is SGAC but not RNIC+DF. RNIC removes $\{(2,3),(3,2)\}$ from

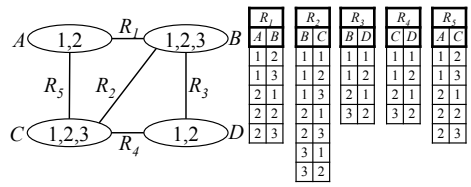

Figure 6: The CSP is SGAC but not RNIC+DF.

R_2, $\{(1,2),(1,3)\}$ from R_1, and $\{(1,2),(1,3)\}$ from R_5. Therefore, RNIC+DF removes the value 1 from A. □

Figure 7 shows the relationships between the domain-filtering properties discussed above.

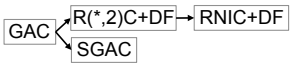

Figure 7: Some domain filtering properties.

3.4 An algorithm for enforcing RNIC

The algorithm for enforcing RNIC is discussed in detail in (Woodward et al. 2011) where it is generically denoted PROCESSQ. Below, we summarize its operation. We define S_τ, the *support* of a tuple $\tau \in R$, to be the set of tuples that verify the condition: $\forall R' \in \text{Neigh}(R), \exists(\tau' \in R'), (\tau' \in S_\tau)$, and the tuples in $S_\tau \cup \{\tau\}$ agree on all shared variables. For each tuple $\tau \in R$ without a valid support, the function SEARCHSUPPORT determines S_τ as the first solution of a backtrack search procedure on the subproblem induced by $\{R\} \cup \text{Neigh}(R)$ and where R is assigned τ. The support is recorded and stays active as long as all its constituents tuples are active. SEARCHSUPPORT uses forward checking and dynamic variable ordering (domain/degree). Two major mechanisms significantly contributed to the success of this search process by improving its running time:

1. *The use of the index-tree data structure* (Karakashian et al. 2010) to determine whether or not two tuples of two adjacent relations in the dual graph are consistent.
2. *The dynamic identification, after each dual variable instantiation, of trees in the dual graph of uninstantiated dual variables* (Woodward et al. 2011). The instantiation of a variable eliminates, from the problem, the variable and the constraints that link it to the uninstantiated variables, potentially breaking cycles and yielding trees. We call those trees *dangles*: They can be floating trees or may be attached to some subgraph. Dangles can be identified efficiently and exploited for early inconsistency detection.

Note that dangle identification is a general mechanism for improving the performance of *any* backtrack search. Obviously, it cannot be used in the algorithm for enforcing GAC or R(∗,2)C (where there is no search). Further, it is not particularly useful in the algorithm for enforcing R$(*, m)$C because the values of m are small in practice.

Let e be the number of constraints in the CSP, k the maximum arity of the constraints, d the maximum domain size, $t = \mathcal{O}(d^k)$ the maximum number of tuples in a relation, and δ the degree of the dual graph. The complexity of the algorithm for enforcing RNIC is $\mathcal{O}(t^{\delta+1}e\delta)$.

3.5 Enforcing RNIC versus R(∗,m)C

The above summarized algorithm and that for enforcing R(∗,m)C (Karakashian et al. 2010) are similar in that they both try to 'complete' (Freuder 1991) each tuple in each relation over one (or more) sets of relations.

The algorithm for R(∗,m)C considers *every* combination of m connected relations. The number of those combinations is $O(e^m)$. Further, each relation needs to be 'checked' against $m-1$ relations in *each* combination where it appears.

The algorithm for enforcing RNIC does not suffer from the above drawbacks. First, the number of combinations considered is equal to the number of relations (e), and each relation is 'checked' against a unique set of relations, which is determined by its neighborhood. Further, the size of the neighborhood is determined *locally* by the connectivity of the relation in the dual graph. Thus, the 'level' of consistency enforced is not necessarily the same on all relations of the dual graph: Lower levels are enforced on sparser portions of the dual graph, and higher levels on the denser portions. In particular, on a cycle of length four or more, RNIC 'naturally' reduces to R(∗,2)C, see Theorem 1.

4 Reformulating the Dual Graph

Two topological conditions of the dual graph can seriously hinder the performance of PROCESSQ:

1. *High density of the dual graph.* As the density of the dual graph increases, the neighborhood of a given relation R_i grows, which increases the cost of enforcing RNIC. To address this issue, we reformulate the dual graph by removing redundant edges.
2. *The existence of cycles of length four or more.* On a cycle of length four or more, the two adjacent relations of a given relation R_i in the cycle are prevented from 'communicating,' thus reducing RNIC to R(∗,2)C (see Theorem 1). To address this issue, we propose to reformulate the dual graph by triangulation,[5] which eliminates cycles of length four or more.

The above two reformulations have the following effects:

- Removing redundant edges cannot strengthen the consistency property enforced by the algorithm and cannot decrease the number of nodes visited by search.
- Adding edges by graph triangulation cannot weaken the consistency property enforced and cannot increase number of nodes visited by search.

Applying PROCESSQ on the dual graph reformulated by one or both of the above reformulations enforces three variations of RNIC, namely wRNIC, triRNIC, and wtriRNIC, where the prefixes 'w' and 'tri' denote the consistency properties resulting from removing redundant edges and triangulating the dual graph, respectively. Figure 8 illustrates those relationships in a partial order. Naturally, the property enforced depends on the particular minimal and triangulated dual graph used.

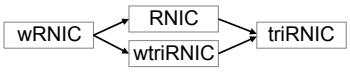

Figure 8: Variations of RNIC.

While the *set of solutions to the CSP is not affected by either reformation*, it is not straightforward to predict the effect of the above reformulations on CPU time. To lay it out, we would like to remove enough edges from the dual graph to reduce the running time of PROCESSQ, which is $\mathcal{O}(t^{\delta+1}e\delta)$. However, we would also like to add enough

[5] Graph triangulation adds an edge (a chord) between two non-adjacent vertices in every cycle of length four or more (Golumbic 2004). While minimizing the number of edges added by the triangulation process is NP-hard, MINFILL is an efficient heuristic commonly used for this purpose (Kjærulff 1990; Dechter 2003).

edges to the dual graph in order to boost propagation. Furthermore, we need a strategy to automatically select the appropriate reformulation. Below, we discuss the two reformulations (Sections 4.1 and 4.2) and their combination (Section 4.3). In Section 4.4, we propose a procedure to automatically select a reformulation in a preprocessing step.

4.1 Removing redundant edges: wRNIC

An edge between two vertices in the dual graph is *redundant* if there exists an alternate path between the two vertices such that the shared variables appear in every vertex in the path (Janssen et al. 1989; Dechter 2003). Redundant edges can be removed without affecting the set of solutions of the CSP. Janssen et al. (1989) introduced an efficient algorithm for computing the *minimal dual graph* by removing redundant edges. Many minimal graphs may exist, but all are guaranteed to have the same number of edges. Figure 9 shows the dual graph (density 60%) and a minimal dual graph (density 40%) of the example of Figure 1. Note that R(∗,2)C

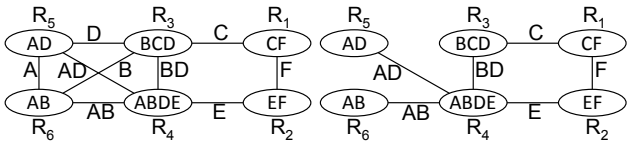

Figure 9: A minimal dual graph.

≡ wR(∗,2)C (Janssen et al. 1989). Also, computing and storing the combinations of relations necessary for enforcing R(∗,m)C is not possible in practice unless the redundant edges are first removed from the dual graph (Karakashian et al. 2010).

Our experiments showed that RNIC is advantageous on dual graphs of density up to around 15%.[6] For higher density values, we propose to remove the redundant edges in the dual graph before running PROCESSQ. This operation reduces the density of the original dual graph and the size of the induced subproblems on which SEARCHSUPPORT is executed. It also results in a weakened consistency, denoted wRNIC, that depends of the particular minimal graph used. Because wRNIC is enforced on a minimal dual graph (i.e., a graph with no more edges than the original dual graph), wRNIC is strictly weaker than wRNIC.

Figure 10 integrates the above discussion in the partial order of Figure 2. Note that these results hold between the

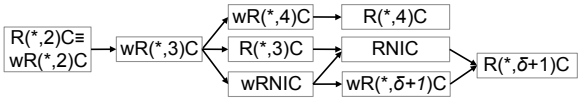

Figure 10: Relating RNIC, wRNIC, R(∗,m)C, and wR(∗,m)C.

weakened properties provided they are enforced on the *same* minimal dual graph.

[6] In a related research, we studied the density of 1689 dual graphs of (binary and non-binary) CSPs from the Solver Competition Benchmarks. We identified a sharp threshold at 15% density. Indeed, 56.6% of the dual graphs (79.9% after redundancy removal) considered had a density less than or equal to 15%. It is not yet clear to us how to interpret the value of this threshold.

4.2 Triangulating the dual graph: triRNIC

When the dual graph has only cycles of size four or more, RNIC reduces to R(∗,2)C (see Theorem 1), which significantly hampers filtering and propagation. To remedy this situation, we propose to triangulate the dual graph. This process creates loops in the dual graph and increases the size of the induced subproblems on which SEARCHSUPPORT is executed, boosting the propagation process, but also raising the consistency level enforced on the CSP. For example, in the dual graph of the example of Figure 1, $\mathrm{Neigh}(R_1)$=$\{R_2, R_3\}$. However, $\mathrm{Neigh}(R_1)$=$\{R_2, R_3, R_4\}$ in the triangulated graph (density 67%) of Figure 11. We denote the resulting consistency

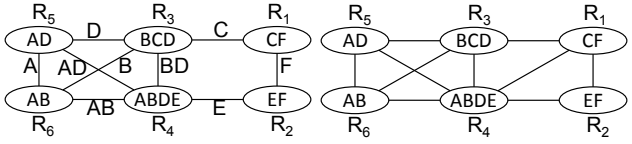

Figure 11: Triangulating a dual graph.

property triRNIC. Similarly to wRNIC, triRNIC depends on the particular triangulation of the dual graph.

An important feature of the triangulation process is that it operates *locally*, adding edges only where cycles of length four or more need to be shortened, irrespective of the degree of the vertices in the graph.

4.3 Triangulate a minimal dual graph: wtriRNIC

While using a minimal dual graph allows us to cope with the high density of difficult benchmark instances, triangulating the minimal dual graph allows us to boost propagation. We denote wtriRNIC the consistency resulting from applying PROCESSQ on the triangulated minimal dual graph. Figure 12 shows the dual graph (density 47%) resulting from applying both reformulations in sequence for the example of Figure 1. As shown in Figure 8, wtriRNIC is strictly stronger

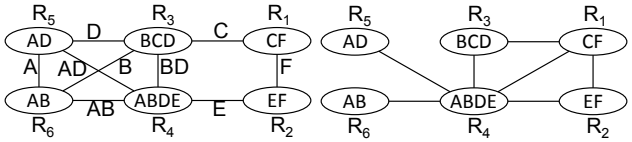

Figure 12: Triangulating a minimal dual graph.

than wRNIC applied on the same minimal dual graph, but strictly weaker than triRNIC. Further, it is not comparable with RNIC, which is enforced on the original dual graph. Figure 13 summarizes the relationships between RNIC, its reformulations, and R(∗,m)C based properties.

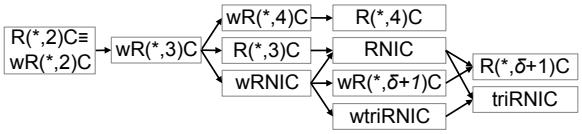

Figure 13: Relating RNIC, R(∗,m)C, and their studied variations.

4.4 Select the appropriate RNIC: selRNIC

The algorithm summarized in Section 3.4, PROCESSQ, enforces any of the four properties RNIC, triRNIC, wRNIC,

and wtriRNIC on a CSP by operating on the original dual graph or some modification of it.

- For RNIC, it uses the original dual graph (G_o).
- For wRNIC, it uses a minimal dual graph (G_w).
- For triRNIC, it uses a triangulated dual graph (G_{tri}).
- Finally, for wtriRNIC, it uses a triangulated minimal dual graph (G_{wtri}).

The selection policy shown in Figure 14 automatically chooses the dual graph on which to enforce RNIC by comparing the density d^G of a given dual graph G. The goal of this deliberation is to adjust the strength of propagation to the topology of the dual graph. Paraphrasing the content

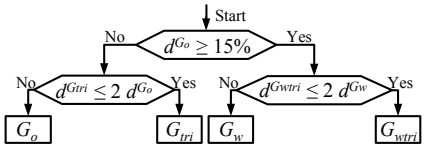

Figure 14: Selecting a dual graph for selRNIC.

of Figure 14, we consider the dual graph of density greater than or equal 15% to be too dense to be effectively processed by PROCESSQ. For this reason, we choose to reformulate it by removing redundant edges. Whenever triangulation does not increase the density of a dual graph more than two fold, then the advantage of boosting propagation by creating loops and increasing neighborhood sizes outweighs the drawback of increasing the cost of operating on larger neighborhoods. For the example of Figure 1, this policy correctly chooses the triangulated minimal dual graph (density 47%). While both operations of triangulating a dual graph and computing a minimal dual graph can be done efficiently and do not add *any* perceptible overhead in our experiments, the policy of Figure 14 applies each operation at most once. The resulting mechanism, which we denote selRNIC, nicely ties together our techniques in a consistent and adaptive framework.

5 Related Work

NIC was proposed by Freuder and Elfe in (1996) and evaluated by them and others on binary CSPs. Debruyne and Bessière (2001) showed that NIC is ineffective on sparse graph and too costly on dense graphs. Below, we restrict our discussion to non-binary CSPs. In (Bacchus et al. 2002), *nic(dual)* denotes applying NIC to the dual encoding of a CSP. As stated in the introduction, it is identical to RNIC. However, the paper does not go beyond stating that *nic(dual)* is strictly stronger than *ac(dual)* (i.e., RNIC is strictly stronger than R($*$,2)C). Otherwise, most of the research on consistency for non-binary CSPs has focused on filtering the variables domains, not the constraints definitions, such as the study of 'variable-based' NIC (Gent, Stergiou, and Walsh 2000; Stergiou 2007).

More generally, relational consistency properties were formalized in (Dechter and van Beek 1997) as *relational m-consistency* and *relational (i, m)-consistency*. Enforcing those properties may require adding constraints to the problem, modifying its topology. As for relation-filtering properties, m-wise consistency was proposed in relational databases (Gyssens 1986). Janssen et al. (1989) showed that arc consistency on the dual encoding of a CSP enforces pairwise consistency. Algorithms for R($*$,m)C, which is equivalent to m-wise consistency, were proposed for arbitrary $m \geq 2$ and evaluated in (Karakashian et al. 2010). One limitation of the algorithm for R($*$,m)C is the need to manually select m and generate all combinations of m relations that form a connected graph. The number of combinations grows exponentially with m, causing space limitations. In comparison, RNIC requires storing for each relation R a unique combination of constraints $\{R\} \cup \mathrm{Neigh}(R)$ and the size of this combination varies with the connectivity of R in the dual graph. Given the space requirement for storing all combinations of m relations, Karakashian et al. (2010) proposed to enforce R($*$,m)C on minimal dual graphs only, namely wR($*$,2)C, wR($*$,3)C, and wR($*$,4)C. The support structures used in PROCESSQ are similar to those proposed in (Bessière et al. 2005). Finally, the insight that breaking cycles yields trees in a search space (i.e., tree, or dangle, identification in SEARCHSUPPORT, Section 3.4) can be related to the Cycle-Cutset method (Dechter and Pearl 1987).

6 Experimental Results

We ran our experiments on the benchmarks of the CSP Solver Competition[7] with a time limit of one and a half hours per instance. Below we report our results, which are quantitatively but not qualitatively different from the results described in (Woodward et al. 2011). We also conduct statistical tests to determine the significance of those results. We have slightly improved the code of all our algorithms and enlarge the memory pool giving benefit to the wR($*$,m)C algorithms. The justification for those changes is to focus more on the performance of the algorithms rather than focusing on their limitations. We report our results first for the entire pool of reported benchmarks (Table 1), then for four representative benchmark problems (Tables 2, 3, and 4). Given that some algorithms did not complete some instances, some data points are missing. For this reason, we consider the data to be right-censored and conduct a survival data analysis (Lee 1992). The survival data analysis does not make any assumption about the distribution of the data, and yields a calculated mean CPU time for each algorithm, reported in column **Time** in Tables 1, 2, 3, and 4. A '-' entry in those columns indicates that, even though the corresponding algorithm terminated on some instances, it did not terminate on enough instances to yield an accurate statistical mean. In column **R**, we report the rank of each algorithm based on the probability of the survival data analysis, breaking ties based on the time for reaching that probability. In column **S**, we group the algorithms into equivalence classes of CPU performance. To compute the statistically significant categories, we perform a pairwise significance check between every two algorithms for a significance level of 0.05. This comparison requires a normal distribution of the non-censored data. For this analysis, we assume that all censored data points finished at the maximum cutoff time. In Table 1, column $\mathbf{S}_B$ provides a coarser classification of those algorithms based on termination only while ignoring CPU time. Regardless of whether or not the normality assumption holds, each analysis yielded similar results, hint-

[7] http://www.cril.univ-artois.fr/CPAI09/

ing that our conclusions are correct. In addition to the above, the tables provide: the number of completed instances (#C); the number of instances with the fastest running time (#F), where ties are awarded to all parties; the number of instances solved backtrack free (#BF); and the range of the density of the dual graph d^D that each of the algorithms works on. Further, for each benchmark class, we report the number of instances in the class and the range of the number of constraints e. Before we discuss the results in detail, we note that the values of nodes visited in all experiments comply with the partial order shown in Figure 13.

Table 1 reports the results of all reported benchmarks. Notice how selRNIC outperforms all other algorithms. It ties only with wR($*$,2)C on the pairwise significance check for CPU time in column **S**. Further, with a 50ms error tolerance, selRNIC outperforms in a statistically significant manner a random selection of the four RNIC-based algorithms.

Table 1: Comparison over the 169 instances reported.

Algorithm	**Time**	**R**	**S**	**S**$_B$	**#C**	**#F**	**#BF**
aim-100, aim-200, lexVg, modifiedRenault, ssa benchmarks							
wR($*$,2)C	944924	3	**A**	B	138	52	79
wR($*$,3)C	925004	4	B	B	134	8	92
wR($*$,4)C	1161261	5	B	B	132	2	108
GAC	1711511	7	C	C	119	**83**	33
RNIC	6161391	8	C	C	100	19	66
triRNIC	3017169	9	C	C	84	9	80
wRNIC	1184844	6	B	B	131	8	84
wtriRNIC	937904	2	B	B	144	3	129
selRNIC	**751586**	**1**	**A**	**A**	**159**	17	**142**

Table 2 illustrates the usefulness of RNIC: it solves the largest number of problems in this set, and solves, backtrack free, the largest number of instances. In terms of significance ranking, GAC, triRNIC, and wRNIC are not advantageous techniques for these problems that have low density, and high density after triangulation. selRNIC was able to select the dual graph that yielded the largest number of completions and backtrack free. Despite not always being the fastest, it was not significantly different than the algorithm that was the fastest.

Table 3 illustrates the usefulness of wRNIC and wtriRNIC. As stated above, the sheer number of relations combined with the large density in the dual graphs of the problems in this benchmark prevents us from executing RNIC and triRNIC. This situation demonstrates the benefits of using wRNIC and wtriRNIC, which were actually automatically chosen by selRNIC. Note also that wtriRNIC solves, backtrack free, all instances in this category. We cannot stress enough on the importance of this last fact: It is indicative of the tractability of this class of problems. Notice, despite selRNIC not having the smallest CPU time, there is not a statistically significant difference between the mean CPU time of selRNIC and and the mean CPU time of wR($*$,2)C. Once again, GAC was in a lower significance class than selRNIC, as with RNIC and triRNIC, as was expected.

In both Tables 2 and 3, selRNIC largely outperforms GAC for all measures. Even if one was to use a high-performance GAC implementation such as the one in (Cheng and Yap 2010), the number of nodes visited by GAC remains orders of magnitude larger than that by selRNIC, and the number of

Table 2: RNIC/selRNIC completes the largest number of instances, and solves, backtrack free, the largest number of instances.

Algorithm	d^D	**Time**	**R**	**S**	**#C**	**#F**	**#BF**
aim-100: 24 instances, $e \in$[150,570]							
wR($*$,2)C	[6.3%,8.1%]	412369	5	**A**	19	6	5
wR($*$,3)C		304816	3	**A**	20	1	7
wR($*$,4)C		140070	2	**A**	20	0	12
GAC	N/A	1923579	7	B	17	4	1
RNIC/ selRNIC	[6.3%,8.1%]	**94699**	**1**	**A**	**22**	5	**16**
triRNIC	[26.0%,70.5%]	2259986	8	B	9	1	9
wRNIC	[0.7%,2.6%]	1009380	4	B	20	**8**	7
wtriRNIC	[6.3%,8.1%]	1280885	6	**A**	17	0	8
aim-200: 24 instances, $e \in$[302,1169]							
wR($*$,2)C	[3.2%,4.2%]	**132205**	5	B	12	**10**	4
wR($*$,3)C		1006472	2	**A**	15	3	8
wR($*$,4)C		2015651	3	B	12	0	8
GAC	N/A	-	6	C	8	0	0
RNIC/ selRNIC	[3.2%,4.2%]	781596	**1**	**A**	**19**	5	**13**
triRNIC	[21.2%,71.6%]	-	8	C	1	0	1
wRNIC	[0.4%, 1.4%]	244643	4	B	13	3	5
wtriRNIC	[6.3%,11.6%]	-	7	C	6	0	6

Table 3: RNIC is hindered by the high density of the dual graph, but its weakened versions outperform all others.

Algorithm	d^D	**Time**	**R**	**S**	**#C**	**#F**	**#BF**
modifiedRenault 50 instances, $e \in$[147,159]							
wR($*$,2)C	[35.4%,41.6%]	**3078**	6	**A**	46	**30**	41
wR($*$,3)C		8463	3	**A**	49	4	48
wR($*$,4)C		31157	**1**	**A**	**50**	2	**50**
GAC	N/A	1678928	7	B	25	14	4
RNIC	[35.4%,41.6%]	-	8	B	7	0	7
triRNIC	[36.4%, 43.8%]	-	9	B	5	0	5
wRNIC	[1.7%,1.9%]	8285	5	**A**	47	0	43
wtriRNIC	[2.9%,3.9%]	166652	2	**A**	**50**	0	**50**
selRNIC	[1.8%,3.6%]	166560	4	**A**	49	0	48

instances solved backtrack-free significantly smaller. Only in Table 4 does GAC outperform the other algorithms in terms of CPU time only. Interestingly, however, on lexVg, and despite the high density ([57.6%,78.6%]) of the redundancy removed triangulated dual graph, wtriRNIC/selRNIC solves in a backtrack-free manner all but one of the instances in this set, thus hinting to the tractability of these instances. (The last instance hit the time threshold.) Notice that even though GAC has a smaller CPU time than selRNIC, the difference between the two algorithms is not statistically significantly (see column **S**). There were not enough instances (8) for the ssa benchmark, reported in (Woodward et al. 2011), to report any statistically significant conclusions.

7 Future Work & Conclusions

Our approach opens the door to the investigation of a new type of singleton consistency properties for non-binary CSPs. Instead of assigning the value of a *single* variable before enforcing some level of consistency on the CSP, as it is usually the case for Singleton Arc Consistency (SAC) (Bessiere et al. 2011), we should investigate the effectiveness of 'assigning a tuple to a relation' in the dual problem. Such an approach would yield a new class of relational consistency properties, which could be called *relation-based singleton consistency properties*. Note however, that, un-

Table 4: GAC is best on CPU, triRNIC/selRNIC is best on #BF.

Algorithm	d^D	Time	R	S	#C	#F	#BF
lexVg: 63 instances, $e \in$[8,36]							
wR(∗,2)C	[48.5%,57.1%]	809765	4	C	55	4	27
wR(∗,3)C		1384983	7	C	44	0	27
wR(∗,4)C		1525548	7	C	28	0	26
GAC	N/A	**114827**	**1**	**A**	**63**	**61**	26
RNIC	[48.5%,57.1%]	1647671	6	C	45	7	27
triRNIC	[57.6%,78.6%]	1031882	3	B	62	7	**62**
wRNIC	[48.5%,57.1%]	1464461	5	C	43	1	27
wtriRNIC/ selRNIC	[57.6%,78.6%]	**580935**	2	**A**	62	7	**62**

like RNIC, maintaining such properties during search is prohibitive in practice (Lecoutre and Prosser 2006).

Our algorithm operates on relations defined in extension as consistent tuples (supports). Relations defined in extension as conflicts (no-goods) could be converted to supports, as we did here. Further, and also for constraints defined in intension, we could generate support tuples after applying GAC to the original CSP. For cases where it is important to keep all relation definitions in intension, we claim that a similar, albeit weaker, domain pruning can be achieved by executing RNIC on combinations of domain values that are consistent with the relations. We propose to mitigate the loss of information by generating new (support) constraints of some judiciously chosen scopes. We propose to investigate this approach in the future and evaluate its effectiveness.

Consistency properties and their algorithms are central to CP, and perhaps best distinguish this discipline from other fields that study the same problems. Research has focused on defining new properties, proposing new algorithms, improving the performance of known ones, and theoretically characterizing the relationship between the consistency level and the tractability of the CSP. Our contribution exploits and adds to the large body of literature on consistency properties and their propagation algorithms. However, our long-term goal is to design techniques that allow a constraint solver to identify tractable problem classes and automatically select and apply the appropriate tools for solving them. In that sense, the ability of our techniques to adapt to a problem's structure and solve many difficult instances in a backtrack-free manner[8] is perhaps the most noteworthy contribution of the current research: It indicates that we may be one step closer to achieving our goal.

Acknowledgments

We are grateful to Elizabeth Claassen and David B. Marx of the Department of Statistics at the University of Nebraska-Lincoln (UNL) for their help with designing the statistical analysis. Experiments were conducted on the equipment of the Holland Computing Center at UNL. Robert Woodward was partially supported by a B.M. Goldwater Scholarship and by a National Science Foundation (NSF) Graduate Research Fellowship grant number 1041000. This research is supported by NSF Grant No. RI-111795.

References

Bacchus, F.; Chen, X.; Beek, P. V.; and Walsh, T. 2002. Binary vs. Non-Binary Constraints. *Artificial Intelligence* 140:1–37.

Bessière, C.; Régin, J.-C.; Yap, R. H.; and Zhang, Y. 2005. An Optimal Coarse-Grained Arc Consistency Algorithm. *Artificial Intelligence* 165(2):165–185.

Bessiere, C.; Cardon, S.; Debruyne, R.; and Lecoutre, C. 2011. Efficient Algorithms for Singleton Arc Consistency. *Constraints* 16 (1):25–53.

Bessière, C.; Stergiou, K.; and Walsh, T. 2008. Domain Filtering Consistencies for Non-Binary Constraints. *Artificial Intelligence* 172:800–822.

Bessiere, C. 2006. *Handbook of Constraint Programming*. Elsevier. chapter Constraint Propagation.

Cheng, K. C., and Yap, R. H. 2010. An MDD-Based Generalized Arc Consistency Algorithm for Positive and Negative Table Constraints and Some Global Constraints. *Constraints* 15 (2):265–304.

Debruyne, R., and Bessière, C. 1997. Some Practicable Filtering Techniques for the Constraint Satisfaction Problem. In *Proc. of the 15 th IJCAI*, 412–417.

Debruyne, R., and Bessière, C. 2001. Domain Filtering Consistencies. *Journal of Artificial Intelligence Research* 14:205–230.

Dechter, R., and Pearl, J. 1987. The Cycle-Cutset Method for improving Search Performance in AI Applications. In *Third IEEE Conference on AI Applications*, 224–230.

Dechter, R., and van Beek, P. 1997. Local and Global Relational Consistency. *Theor. Comput. Sci.* 173(1):283–308.

Dechter, R. 2003. *Constraint Processing*. Morgan Kaufmann.

Freuder, E. C., and Elfe, C. D. 1996. Neighborhood Inverse Consistency Preprocessing. In *Proc. of AAAI-96*, 202–208.

Freuder, E. C. 1991. Completable Representations of Constraint Satisfaction Problems. In *Second International Conference on Principles of Knowledge Representation and Reasoning*, 186–195.

Gent, I.; Stergiou, K.; and Walsh, T. 2000. Decomposable Constraints. *Artificial Intelligence* 123 (1-2):133–156.

Golumbic, M. C. 2004. *Algorithmic Graph Theory and Perfect Graphs*. Elsevier. Annals of Discrete Mathematics, Vol 75.

Gyssens, M. 1986. On the Complexity of Join Dependencies. *ACM Trans. Database Systems* 11(1):81–108.

Janssen, P.; Jégou, P.; Nougier, B.; and Vilarem, M. 1989. A Filtering Process for General Constraint-Satisfaction Problems: Achieving Pairwise-Consistency Using an Associated Binary Representation. In *IEEE Workshop on Tools for AI*, 420–427.

Karakashian, S.; Woodward, R.; Reeson, C.; Choueiry, B. Y.; and Bessiere, C. 2010. A First Practical Algorithm for High Levels of Relational Consistency. In *AAAI 10*, 101–107.

Kjærulff, U. 1990. Triagulation of Graphs - Algorithms Giving Small Total State Space. Research Report R-90-09, Aalborg University, Denmark.

Lecoutre, C., and Prosser, P. 2006. Maintaining Singleton Arc Consistency. In *CPAI 06 Workshop on Symmetry in Constraint Satisfaction Problems (SymCon 10)*, 47–61.

Lee, E. T. 1992. *Statistical Methods for Survival Data Analysis*. New York, NY: John Wiley & Sons, second edition.

Stergiou, K. 2007. Strong Inverse Consistencies for Non-Binary CSPs. In *Proc. of the 19th IEEE International Conference on Tools with Artificial Intelligence*, volume 1 of *ICTAI 07*, 215–222.

Woodward, R.; Karakashian, S.; Choueiry, B. Y.; and Bessiere, C. 2011. Solving Difficult CSPs with Relational Neighborhood Inverse Consistency. In *AAAI 11*, 1–8.

[8]Note that the complexity of RNIC is exponential in the degree of the dual graph and not in the number of variables.