Full Constraint Satisfaction Problems

Tomás Feder
268 Waverley St., Palo Alto, CA 94301, USA
tomas@theory.stanford.edu,
Pavol Hell
School of Computing Science
Simon Fraser University
Burnaby, B.C., Canada V5A 1S6
pavol@cs.sfu.ca

Abstract

Feder and Vardi have conjectured that all constraint satisfaction problems to a fixed structure (constraint language) are polynomial or NP-complete. This so-called Dichotomy Conjecture remains open, although it has been proved in a number of special cases. Most recently, Bulatov has verified the conjecture for conservative structures, i.e., structures which contain all possible unary relations.

We explore three different implications of Bulatov's result. Firstly, the above dichotomy can be extended to so-called inclusive structures, corresponding to conservative constraint satisfaction problems in which each variable comes with its own domain. (This has also been independently observed by Bulatov.) We prove a more general version, extending the dichotomy to so-called three-inclusive structures, i.e., structures which contain, with any unary relation R, all unary relations R' for subsets $R' \subseteq R$ with at most three elements.

For the constraint satisfaction problems in this generalization we must restrict the instances to so-called 1-full structures, in which each variable is involved in a unary constraint. This leads to our second focus, which is on restrictions to more general kinds of 'full' input structures. For any set W of positive integers, we consider a restriction to W-full input structures, i.e., structures in which, for each $w \in W$, any w variables are involved in a w-ary constraint. We identify a class of structures (the so-called W-set-full structures) for which the restriction to W-full input structures does not change the complexity of the constraint satisfaction problem, and hence the family of these restricted problems also exhibits dichotomy. The general family of three-inclusive constraint satisfaction problems restricted to W-full input structures contains examples which we cannot seem to prove either polynomial or NP-complete. Nevertheless, we are able to use our result on the dichotomy for three-inclusive constraint satisfaction problems, to deduce the fact that all three-inclusive constraint satisfaction problems restricted to W-full input structures are NP-complete or 'quasi-polynomial' (of order $n^{O(\log n)}$).

Our third focus deals with bounding the number of occurrences of a variable, which we call the degree. We conjecture that the complexity classification of three-inclusive constraint satisfaction problems extends to the case where all degrees are bounded by three. Using previous results, we are able to verify this conjecture in a number of special cases.

Conservative, inclusive, and three-inclusive constraint satisfaction problems can be viewed as problems in which each variable is restricted to a 'list' of allowed values. This point of view of lists is frequently encountered in the study of graph colorings, graph homomorphisms, and graph partitions. Our results presented here, in all three areas, were strongly motivated by these results on graphs.

1 Introduction

A large class of problems in artificial intelligence and other areas of computer science can be viewed as *constraint-satisfaction problems* [8, 31, 32, 33, 34, 40]. This includes problems in machine vision, belief maintenance, scheduling, temporal reasoning, type reconstruction, graph theory, and satisfiability.

The standard formulation of the constraint satisfaction problem goes back to Montanari [36] in 1974. This framework has proved its value over the years by its wide ranging applicability [8, 39]. The constraint satisfaction in its full generality is NP-complete. Thus constraint satisfaction problems have been studied under various restrictions. The main model [20] considers constraint satisfaction problems with a fixed template determining the size of the domain and the set of allowed constraint types in an instance.

Formally, a *vocabulary* V is a set of pairs (R_i, ℓ_i), where the R_i are *relation names* and the ℓ_i are *relation arities*. A *structure* S *over the vocabulary* V consists of a set (called the *domain*) D, together with a collection of relations $R_i \subseteq D^{\ell_i}$, one ℓ_i-ary relation R_i for each pair (R_i, ℓ_i) in V. We say that the structure S an *interpretation* of the vocabulary V, and emphasize this, if necessary, by writing V^S instead of S. We also write the domains and relations as D^S and R_i^S, to indicate the structure in which they are being interpreted.

The *constraint satisfaction problem* $CSP(H)$ (or $CSP(V^H)$) for a fixed structure $H = V^H$ over a vocabulary V is given as follows. An *instance* of $CSP(H)$ is a structure V^G over the same vocabulary V. The *question* asked by $CSP(H)$ is whether or not there exists a *homomorphism* f of G to H, that is, a mapping $f : D^G \to D^H$ such that $(x_1, \ldots, x_{\ell_i}) \in R_i^G$, then $(f(x_1), \ldots, f(x_{\ell_i})) \in R_i^H$, for all choices of $x_i \in D^G$ and $(R_i, \ell_i) \in V$. We shall speak of the elements x_i of D^G as *variables*, constrained by the relations R_i^G to be assigned suitable *values* $f(x_i)$ in D^H, as allowed by the constraints R_i^H.

This model was studied in its full generality by Feder and Vardi [20], who found evidence towards the following *Dichotomy Conjecture:*

Conjecture 1.1 *For each fixed structure H, the problem $CSP(H)$ is NP-complete or polynomial time solvable.*

(Recall that it is known that if P $\neq$ NP, there are problems in NP that are neither NP-complete, nor polynomial [30].) This conjecture was supported by earlier classifications that exhibited dichotomies. Schaefer [38] classified the complexity of Boolean constraint satisfaction problems, i.e., problems $CSP(H)$ with a two-element domain $D^H = \{0, 1\}$. In essence, the polynomial cases consist of Horn or anti-Horn clauses, two-satisfiability, and linear equations modulo two. The remaining constraint satisfaction problems $CSP(H)$ for a two-element domain D^H are all NP-complete. Bulatov classified constraint satisfaction problems on a three-element domain [3]. Hell and Nešetřil [26] classified the complexity of constraint satisfaction problems $CSP(H)$ where $H = V^H$ is a graph (with loops allowed). In other words, V consists of one pair $(R, 2)$, and the relation R^H is symmetric. They showed that this so-called *H-coloring problem* is polynomial if H is bipartite, and is NP-complete for H non-bipartite. Such a result is not known for digraphs (structures V^H in which R^H is not necessarily symmetric) [1, 20]. In fact, Feder and Vardi [20] have shown that dichotomy for digraph H-coloring problems would imply the validity of the entire Dichotomy Conjecture.

For this paper, the most relevant special case where the conjecture has been proved is the following result. We say that a structure H is *conservative* if it contains a unary relation R for each subset R of the domain D^H. We shall denote the set of unary relation names (names R

involved in pairs $(R,1)$) of a vocabulary V by $\mathcal{U}(V)$, and also denote the set of the corresponding relations R^H in a structure $H = V^H$ by $\mathcal{U}(V^H)$ (or just $\mathcal{U}(H)$). (Thus the vocabulary V of a conservative structure V^H with $|D^H| = n$ has $|\mathcal{U}(V)| = 2^n$, and $\mathcal{U}(V^H)$ is precisely the power set of D^H.) Bulatov [2] proved that dichotomy holds for all conservative structures, i.e., that *for every conservative structure H, the problem $CSP(H)$ is NP-complete or polynomial time solvable.* His approach is based on tools and techniques of universal algebra [2, 3, 4, 5, 28, 29]. Problems $CSP(H)$, for conservative structures H, will be called *conservative constraint satisfaction problems.*

An alternative interpretation of Bulatov's result is the following. Suppose H is a conservative structure; we may restrict $CSP(H)$ to those input structures G that contain each variable $v \in D^G$ in some unary relation. Such structures G will be called 1-*full.* The restriction of $CSP(H)$ to 1-full instances G will be denoted by $CSP_1(H)$. Since H is conservative this is not a real restriction, as there exists in $\mathcal{U}(V)$ a pair $(R,1)$ with the corresponding unary relation R^H equal to D^H. Thus the relation R^G can be imposed an any variable in D^G without changing the problem, and so $CSP(H)$ and $CSP_1(H)$ are equivalent problems. On the other hand, $CSP_1(V^H)$ is easily seen equivalent to the following *list constraint satisfaction problem.* The *instance* is a structure V^G in which each variable $v \in D^G$ is equipped with a list $L(v) \subseteq D^H$, and the *question* is whether or not there exists a homomorphism f of G to H such that each $f(v)$ belongs to $L(v)$. List constraint satisfaction problems have been extensively studied in the case when H is a graph. The problem is polynomial if H is a so-called bi-arc graph and is NP-complete otherwise [12, 13, 14, 15, 16]. Let C be a fixed circle with two specified points p and q. A graph H (with loops allowed) is called a *bi-arc graph* if there exist pairs of arcs $(N_x, S_x), x \in V(H)$, with each N_x being an arc on C containing p but not q, and each S_x being an arc on C containing q but not p, so that for any $x, y \in V(H)$, not necessarily distinct, the following hold:

- either N_x intersects S_y and N_y intersects S_x, or N_x does not intersect S_y and N_y does not intersect S_x;
- N_x intersects S_y (and N_y intersects S_x) if and only if x and y are not adjacent in H.

This class of graphs conveniently generalizes both the class of interval graphs, and the class of (complements of) circular arc graphs of clique covering number two [14, 22].

List constraint satisfaction problems involving multiple binary relations were studied by Feder, Madelaine, and Stewart [19].

Let us say that a structure H is *inclusive* if for any $R \in \mathcal{U}(H)$ and any $R' \subseteq R$ we also have $R' \in \mathcal{U}(H)$. Note that each conservative structure is inclusive. It was independently observed by the present authors and by Bulatov, that the above result for conservative structures H extends to all inclusive structures, in the following sense. *For every inclusive structure H, the problem $CSP_1(H)$ is NP-complete or polynomial time solvable.* Since for conservative structures H the problems $CSP_1(H)$ and $CSP(H)$ are equivalent, this extends the original result of Bulatov [2]. It is easy to interpret this result as the dichotomy of a list constraint satisfaction problem where the variables range over multiple domains. In this context, Feder has previously classified a family of list problems involving multiple Boolean domains [10].

We prove in this paper a more general version of the dichotomy for inclusive structures. We say that a structure H is *three-inclusive* if for any $R \in \mathcal{U}(H)$ and all $R' \subseteq R$ with $|R'| \leq 3$ we also have $R' \in \mathcal{U}(H)$. Thus each inclusive (and hence each conservative) structure is also three-inclusive. We shall prove that *for any three-inclusive structure H, the problem $CSP_1(H)$ is NP-complete or polynomial time solvable.* Problems $CSP_1(H)$, for three-inclusive structures H, will be called *three-inclusive constraint satisfaction problems.*

Three-inclusive structures H are the main focus of this paper. Note that a structure H without unary relations is three-inclusive (in fact inclusive), and thus proving dichotomy for $CSP(H)$ would yield dichotomy for digraph H-coloring problems, and hence the entire dichotomy conjecture. Therefore we restrict the instances G as well. The restriction $CSP_1(H)$ mentioned above is one natural way - restricting to 1-full instances G amounts to assuming the inputs are equipped with lists, as described above. Thus requiring H to be three-inclusive and G to be 1-full gives the three-inclusive constraint satisfaction problem $CSP_1(H)$ that is the connecting thread of our paper. (We then consider further restrictions on both G and H.)

This suggests looking at other versions of 'fullness'. Let W be a set of positive integers such that $1 \in W$, and let V be a vocabulary which contains at least one pair (R, w) for each $w \in W$. We shall say that a structure $G = V^G$ is *W-full* if for each $w \in W$, and any distinct variables $x_1, x_2, \ldots, x_w$ in D^G there exists a permutation σ of $\{1, 2, \ldots, w\}$ and a pair (R, w) in V, such that $(x_{\sigma(1)}, \ldots, x_{\sigma(k)}) \in R^G$. Note that the notion of a 1-full structure coincides with the notion of W-full structure with $W = \{1\}$. We shall also focus on the next natural special case when $W = \{1, 2\}$; in this case we simply say that a W-full structure is *pairwise full* (remembering that, despite the suggestive name, it is assumed to be at the same time 1-full and 2-full). The problem $CSP_W(H)$ is the restriction of $CSP(H)$ to W-full instances G.

A W-full structure $G = V^G$ is *strictly W-full*, if the above permutation σ and relation R are unique, for each $w \in W, w \geq 2$, and any distinct variables $x_1, x_2, \ldots, x_w$ in D^G. The restriction of $CSP(H)$ to strictly W-full instances G, will be denoted by $CSP^*_W(H)$. The notion of *strictly pairwise full* structure refers to, as would be expected, a strictly W-full structure with $W = \{1, 2\}$. A graph G (with loops allowed) can be viewed as strictly pairwise full structure with two binary relations, corresponding to the edges and the nonedges of G. (To conform to the definition as given, we choose an arbitrary orientation of G, i.e., consider each edge uv of G as an oriented arc – either $\vec{uv}$ or $\vec{vu}$ – and similarly for each nonedge.) Under this interpretation, the problems $CSP^*_{\{1,2\}}(H)$ have been studied as *list matrix partition problems* [6, 18, 17, 23, 27]. (The structures H are usually taken to be conservative structures with two binary relations whose union is $(D^H)^2$.)

After Jeavons' work [28, 29] it is usual, when studying the complexity of $CSP(H)$, to consider the relations 'definable in H' along with the relations of H. The properties of structures considered in this paper are sort of fragile with respect to adding definable relations. Thus to convert a $CSP(H)$ problem to a full problem $CSP_W(H)$ one needs to add to H relations of the arities in W that hold for all tuples of elements in D. To convert a full problem $CSP_W(H)$ further to a strict problem $CSP^*_W(H)$ one needs to further add to H relations of each arity in W corresponding to intersection of relations of this arity in H.

This family of (strictly) full problems is not directly under the scope of the Dichotomy Conjecture, and many problems in this class resist attempts at classification [6, 18, 23]. Consider, for instance, the following problem: the vocabulary V consists of eight unary relation names U_i (in other words, V contains eight pairs $(U_i, 1)$), and three binary relation names R_0, R_1, R_2 (additional pairs $(R_0, 2), (R_1, 2), (R_2, 2)$ in V). Now consider the structure V^H with domain $D = D^H = \{0, 1, 2\}$, where the interpretations of the three binary relations are $R^H_j = D^2 - (j, j)$, for $j = 0, 1, 2$, and the eight unary relations correspond to the eight subsets of D. (Assume that U_7 is the relation name for which $U^H_7 = D^H$.) The complexity of $CSP^*_{\{1,2\}}(H)$ is not known. It will follow from Theorem 3.2 that the problem can be solved in quasi-polynomial time $n^{O(\log n)}$ (where $n = |D^G|$). We first illustrate the idea of the recursive algorithm from the proof on Theorem 3.2. For simplicity, we shall assume that the instance V^G has all variables only in the unary relation U^G_7 (having the lists D^H) which does not restrict them in any way. (It is easy to argue that this version is equivalent to the original problem.) Thus we want to solve the following combinatorial problem.

Given a complete graph G with edges colored $0, 1, 2$, can the vertices also be colored by $0, 1, 2$ so that there is no monochromatic edge (an edge of color i whose both endpoints are also colored i) ?

For each vertex x of G we choose a *majority color* $m(x)$ from $0, 1, 2$, with the property that at least a third of the edges incident to x have color $m(x)$. Suppose G has n vertices. We reduce the problem for G to $n + 1$ subproblems as follows. In the first subproblem we assume no vertex x obtains its majority color. In the remaining n subproblems we assume, for each vertex x in turn, that (at least) x obtains its majority color. The first subproblem leaves us with two choices of color for each vertex, and all constraints expressible as clauses with two variables - and hence can be solved by a two-satisfiability algorithm. Each of the other n problems yields a vertex that has been colored, and hence can be removed from the graph; moreover, it implies, for all vertices y adjacent to x along edges of the color $m(x)$, that $m(x)$ is not a legal choice of color for y. In the general step, we have some vertices with lists (of allowed values) of size two and some with lists of size three; we define the majority color $m(x)$ to be the color of at least one third of the edges from x to vertices with lists of size three. Therefore the time to solve a problem with p vertices with lists of size three is

$$T(p) \leq (1 + pT(2p/3)) \cdot T_2(n),$$

where $T_2(n)$ is the (polynomial) time to solve the two-satisfiablity problem arising from n vertices. The recurrence implies that $T(n)$ is quasi-poynomial $n^{O(\log n)}$. Despite the simplicity of this algorithm we have not been able to find a polynomial time algorithm. We have, however, obtained an algorithm of time complexity $n^{O(\log n/ \log\log n)}$ [18].

The analogous problem with colors $0, 1$ corresponds to the well-known graph problem of recognition of *split graphs*, which has a simple polynomial time algorithm [22] (cf. also [24, 27]). When there are more than three colors, the problem is easily seen to be NP-complete. (Our problem can be shown to be no easier than the so-called *stubborn problem* from [6], which also resists classification.)

We shall conclude, using the dichotomy of $CSP(H)$ for three-inclusive structures, that *for a three-inclusive structure H, each problem $CSP_W(H)$ (respectively $CSP^*_W(H)$) is NP-complete or solvable in quasi-polynomial time $n^{O(\log n)}$*. Thus we obtain a new kind of dichotomy - not between problems that are NP-complete and those that are solvable in polynomial time, but between problems that are NP-complete and those that are solvable in quasi-polynomial time. For convenience and brevity we shall call such a result a *quasi-dichotomy*. Our interest in quasi-dichotomy is motivated by arguments similar to those for dichotomy. It is generally expected that no NP-complete problem admits a quasi-polynomial algorithm, and if one NP-complete problem admitted such an algorithm then so would all others. The relationship between the dichotomy conjecture and quasi-dichotomies was previously studied in the case of list matrix partition problems by Feder, Hell, Klein and Motwani [17].

Megiddo and Vishkin [35] studied a problem on tournaments that has a quasi-polynomial algorithm. The problem asks for a minimum size dominating set in a tournament on n vertices. A minimum size dominating set always has at most $\log n$ vertices, so the problem can be solved in time $n^{O(\log n)}$. Megiddo and Vishkin showed that the problem of solving a satisfiability instance with $\log^2 n$ variables and n clauses reduces to the dominating set problem in tournaments. Papadimitriou and Yannakakis [37] showed that the dominating set problem for tournaments is complete for the class $LOGSNP$, with the containment of classes $LOGSNP \subseteq LOGNP$, and both classes reduce in polynomial time to problems in $NP[\log^2 n] \cap DSPACE(\log^2 n)$. Here $NP[\log^2 n]$ is the class of problems that can be solved in nondeterministic polynomial time with only $O(\log^2 n)$ bits of nondeterminism.

The classes $LOGNP$ and $LOGSNP$ are classes of problems in NP involving guessing only

$\log n$ elements in an n-element input structure I. More formally, the classes are defined by setting $LOGNP_0$ to be defined by a formula $\{I : \exists S \in [n]^{\log n} \forall x \in [n]^p \exists y \in [n]^q \forall j \in [\log n] \phi(I, s_j, x, y, j)\}$ and setting $LOGSNP_0$ to be defined by a formula $\{I : \exists S \in [n]^{\log n} \forall x \in [n]^p \exists j \in [\log n] \phi(I, s_j, x, j)\}$, where the S in $LOGSNP_0$ means *strict*, I is an input relation, x, y are tuples of first order variables ranging over $[n] = \{1, \ldots, n\}$, and ϕ is a quantifier-free first order expression. We then define $LOGNP$ to be the class of all languages that can be polynomially reduced to a problem in $LOGNP_0$, and similarly for $LOGSNP$ and $LOGSNP_0$.

All our quasi-polynomial algorithms involve guessing $O(\log n)$ vertices and $O(\log n)$ additional bits of information, thus with $O(\log^2 n)$ bits of nondeterminism, and then solving a resulting problem in deterministic polynomial time. Our quasi-polynomial problems are thus in $NP[\log^2 n]$. However, some of these problems are complete for P, and are thus unlikely to be in $DSPACE(\log^2 n)$.

We shall also observe that constraint satisfaction problems for three-inclusive H show an interesting structure when the degree of the variables, i.e., the number of occurrences of each variable, is bounded. With maximum degree two, Boolean constraint satisfaction can express matching and delta-matroid intersection [7, 9, 11], while with maximum degree three it is the same as with unbounded degree. Similar results on degree restrictions for the case of graphs can be found in [15, 21, 25].

We conjecture that every three-inclusive list constraint satisfaction problem that is NP-complete remains NP-complete when restricted to degree three instances. Using existing results we verify this conjecture in four particular families of three-inclusive list constraint satisfaction problems.

A structure V^H is *W-set-full*, if it is three-inclusive, and if for any $w \in W$ and any (not necessarily distinct) pairs $(S_1, 1), \ldots, (S_w, 1)$ in V such that each $|S_i^H| \leq 3$ for $1 \leq i \leq w$, there exists a pair (R, w) in in V and a permutation σ on $\{1, \ldots, w\}$, such that $S^H_{\sigma(1)} \times \cdots \times S^H_{\sigma(w)} \subseteq R^H$. These W-set-full structures H define *W-set-full constraint satisfaction problems* $CSP_W(H)$ and $CSP^*_W(H)$, which will turn out to be equivalent to the corresponding constraint satisfaction problem $CSP_1(H)$ and hence exhibit dichotomy just like the three-inclusive constraint satisfaction problems. When $W = \{1, 2\}$, we call a W-set-full structure (and problem) *pairwise-set-full*.

The *degree of x in a tuple* $t = (t_1, \ldots, t_k)$ is the number of t_i that are equal to x. The *degree of x in a relation R_i* is the sum of the degrees of x in the tuples in R_i. The *degree of x in a structure G* is the sum of the degrees of x in the relations R_i of arity $\ell_i \geq 2$ in G. The *degree of a structure G* is the maximum degree of the elements of D^G in G.

2 Three-Inclusive Constraint Satisfaction Problems

A *polymorphism* of a structure H with domain $D = D^H$ is a mapping $g : D^k \to D$ such that for all relations R_i^H and all tuples $(x_{1j}, \ldots, x_{\ell_i j}) \in R_i^H$, for $1 \leq j \leq k$, we have $(y_1, \ldots, y_{\ell_i}) \in R_i^H$ for $y_t = g(x_{t1}, \ldots, x_{tk})$. Jeavons [28] showed that the complexity of a constraint satisfaction problem $CSP(H)$ is characterized, up to polynomial time reductions, by the complete set of polymorphisms that H has.

Bulatov [2] classified conservative constraint satisfaction problems $CSP(H)$ as polynomial or NP-complete. The polynomial cases are characterized by the existence of three polymorphisms of H, namely $g_1 : D^2 \to D$, $g_2 : D^3 \to D$, and $g_3 : D^3 \to D$, satisfying the following properties:
(1) g_1, g_2, g_3 are *conservative*, that is $g_1(x, y) \in \{x, y\}$, $g_2(x, y, z) \in \{x, y, z\}$, and $g_3(x, y, z) \in \{x, y, z\}$, for all $x, y, z \in D$.

For each $a, b \in D$ with $a \neq b$, at least one of the following (2)-(4) holds:
(2) g_1 is *commutative* on $\{a, b\}$, that is, $g_1(a, b) = g_1(b, a) \in \{a, b\}$;

(3) g_2 is *majority* on $\{a,b\}$, that is, $g_2(a,a,a) = g_2(a,a,b) = g_2(a,b,a) = g_2(b,a,a) = a$ and $g_2(b,b,b) = g_2(b,b,a) = g_2(b,a,b) = g_2(a,b,b) = b$; and
(4) g_3 is *minority* on $\{a,b\}$, that is, $g_3(a,a,a) = g_3(a,b,b) = g_3(b,a,b) = g_3(b,b,a) = a$ and $g_3(b,b,b) = g_3(b,a,a) = g_3(a,b,a) = g_3(a,a,b) = b$.

Using this classification, we can now prove dichotomy of three-inclusive constraint satisfaction problems $CSP_1(H)$.

Theorem 2.1 *If H is three-inclusive, then the constraint satisfaction problem $CSP_1(H)$ is polynomial time solvable or NP-complete.*

Proof. Let H be a three-inclusive structure; we define three auxiliary structures H^*, H', H'' on the same domain $D^{H^*} = D^{H'} = D^{H''} = D^H = D$, as follows.

- The structure H^* is obtained from H by replacing each non-unary relation R_i of H by the relations $R_i^a = R_i \cap (T_1 \times \cdots \times T_{\ell_i})$, for each choice a of unary relations $T_1, \ldots, T_{\ell_i} \in \mathcal{U}(H)$.
- The structure H' is obtained from H by removing all unary relations $S \in \mathcal{U}(H)$ with $|S| > 3$, and replacing each non-unary relation R_i of H by the relations R_i^a, for each choice a of unary relations of the new structure H' (thus all $|T_i| \leq 3$).
- The structure H'' is obtained from H' by adding all missing unary relations S, so that $\mathcal{U}(H'')$ is the power set of D.

We note that all these structures are three-inclusive, and that H'' is conservative. (Also observe that the three structures are over different vocabularies.)

We first prove that $CSP_1(H)$ and $CSP(H^*)$ are polynomially equivalent problems. To reduce $CSP_1(H)$ to $CSP(H^*)$, we transform an instance G of $CSP_1(H)$ to an instance G^* of $CSP(H^*)$; the structure G^* is obtained from G in a similar way as H^* is obtained from H. (Each non-unary relation R_i is replaced by the relations R_i^a.) It is easy to see that a homomorphism of G to H is also a homomorphism of G^* to H^*. On the other hand, it is also the case that a homomorphism of G^* to H^* is a homomorphism of G to H, since G is 1-full. Thus G has a solution in $CSP_1(H)$ if and only if G^* has a solution in $CSP(H^*)$. To reduce $CSP(H^*)$ to $CSP_1(H)$, we transform an instance G of $CSP(H^*)$ to an instance G' of $CSP_1(H)$ obtained from G as follows. If a variable x is not constrained in G (does not appear in any relation), we remove it. For each $(x_1, x_2, \ldots, x_{\ell_i}) \in R_i^a$ in G, we impose the constraint $(x_1, x_2, \ldots, x_{\ell_i}) \in R_i$ as well as the unary constraints $x_i \in T_i$, $i = 1, \ldots, \ell_i$, in G'. It follows that G' is 1-full, i.e., an instance of $CSP_1(H)$; it is again easy to check that G has a solution in $CSP(H^*)$ if and only if G' has a solution in $CSP_1(H)$.

Since H'' is conservative, we know by [2] that the problem $CSP(H'')$ is polynomial time solvable (if there are polymorphisms g_1, g_2, g_3 of H'' as described above) or NP-complete. If $CSP(H'')$ is polynomial, then the three polymorphisms g_1, g_2, g_3 of H'' are also polymorphisms of H', as all relations in H' are in H''. (In particular, $CSP(H')$ is also polynomial.) Moreover, in this case, g_1, g_2, g_3 are also polymorphisms of H^*, and hence $CSP(H^*)$ is also polynomial. Indeed, consider a relation R_i^a of H^*, where a is the choice of unary relations $T_1, T_2, \ldots, T_{\ell_i} \in \mathcal{U}(H)$. Suppose $x = (x_1, \ldots, x_{\ell_i})$, $y = (y_1, \ldots, y_{\ell_i})$, and $z = (z_1, \ldots, z_{\ell_i})$ are in R_i^a. Let a' be the choice of unary relations $T'_1, T'_2, \ldots, T'_{\ell_i} \in \mathcal{U}(H)$ where each $T'_i = \{x_i, y_i, z_i\}$. Since each $T'_i \subseteq T_i$ and H is three-inclusive, each $T'_i \in \mathcal{U}(H)$ and hence also $T'_i \in \mathcal{U}(H')$. Thus x, y, z are also in $R_i^{a'}$, and, since g_1, g_2, g_3 are polymorphisms of H', we must also have $(g_1(x_1, y_1), \ldots, g_1(x_{\ell_i}, y_{\ell_i})) \in R_i^{a'}$, $(g_2(x_1, y_1, z_1), \ldots, g_2(x_{\ell_i}, y_{\ell_i}, z_{\ell_i})) \in R_i^{a'}$, and $(g_3(x_1, y_1, z_1), \ldots, g_3(x_{\ell_i}, y_{\ell_i}, z_{\ell_i})) \in R_i^{a'}$. Since $R_i^{a'} \subseteq R_i^a$, g_1, g_2, g_3 are also polymorphisms of H^*, whence $CSP(H^*)$ (and $CSP_1(H)$) is polynomial.

Suppose instead $CSP(H'')$ is NP-complete. We will reduce $CSP(H'')$ to $CSP(H')$ and then to $CSP(H^*)$, proving that $CSP(H^*)$ (and $CSP_1(H)$) is also NP-complete. Given an instance G of $CSP(H'')$, we may assume that each variable x is constrained by some non-unary relation of G. (If a variable x is only in unary relations of G, then either the intersection of all the subsets of H corresponding to these unary relations is empty, in which case there is no solution, or the intersection is nonempty, and assigning any value from the intersection to x allows us to remove x from the consideration.) Thus each variable x of G occurs in some jth position in some R_i^a in H', where a is the choice of some unary relations $T_1, T_2, \ldots T_{\ell_i}$ with $|T_j| \leq 3$. We form G_1 from G by replacing each unary relation T containing x which is in H'' but not in H' by the unary relation $T' = T \cap T_j$ for x. Note that T' is a relation of H', since $T' \subseteq T_j$. It is easy to see that G_1 has a solution for $CSP(H')$ if and only if G has a solution for $CSP(H'')$. Therefore $CSP(H')$ is NP-complete. The final reduction from $CSP(H')$ to $CSP(H^*)$ is trivial, as each instance G of $CSP(H')$ is also an instance of $CSP(H^*)$. Therefore $CSP(H^*)$ (and $CSP_1(H)$) is also NP-complete. □

Let us call a structure H *sub-unary* if each non-unary relation R_i of H is included in some product $T_1 \times T_2 \times \cdots \times T_{\ell_i}$ of unary relations $T_1, T_2, \ldots, T_{\ell_i} \in \mathcal{U}(H)$. The structure H^* in the above proof is sub-unary, and this was the only property used in the proof. Thus we have also proved the following fact.

Corollary 2.2 *If H is a sub-unary three-inclusive structure, then $CSP(H)$ is polynomial or NP-complete.*

The above proof also shows that two structures which only differ in their unary relations S with $|S| > 3$ have the same behaviour for CSP_1.

Corollary 2.3 *If H and K are three-inclusive structures on the same domain $D^H = D^K$, which have the same non-unary relations and the same unary relations S with $|S| \leq 3$, then $CSP_1(H)$ and $CSP_1(K)$ are both NP-complete or both polynomial.*

Proof. In the proof of Theorem 2.1, the problems $CSP_1(H)$ and $CSP_1(K)$ yield the same derived problems $CSP(H') = CSP(K')$ and $CSP(H'') = CSP(K'')$. □

We can apply the theorem to the following *three-inclusive list H-coloring problem*, where H is a fixed graph (with loops allowed). *An instance* is a graph G, and with each vertex v a list $L(v)$ of vertices of H, such that $|L(v)| \leq 3$. The *question* is whether or not there exists a homomorphism f of G to H with each $f(v) \in L(v)$.

Corollary 2.4 *[15] Each three-inclusive list H-coloring problem is NP-complete or polynomial time solvable.*

In fact, in [15] we classify exactly which three-inclusive list H-coloring problems are polynomial - they turn out to be the same bi-arc graphs defined above.

3 Restriction to Pairwise Full (and Strictly Pairwise Full) Instances

Our general goal is to study the restriction of three-inclusive constraint satisfaction problems to W-full instances. For simplicity we shall first state and prove our results in the special case where

$W = \{1, 2\}$, i.e., for pairwise full instances G. The general case will be treated in the next section, mostly just by pointing out the extra effort required.

We begin by observing that each three-inclusive constraint satisfaction problem $CSP_1(H)$ can be viewed as the restricted problem $CSP_{\{1,2\}}(H')$, where H' is obtained from H by adding the binary relation $D^H \times D^H$ (and the binary relation $D^G \times D^G$ to any instance G, without changing the problem). We prove a strong converse to this observation; in particular, this will show that $CSP_{\{1,2\}}(H)$ and $CSP^*_{\{1,2\}}(H)$ enjoy dichotomy for pairwise-set-full structures H.

Theorem 3.1 *Let H be a pairwise-set-full structure.*

The problems $CSP_{\{1,2\}}(H), CSP^*_{\{1,2\}}(H)$, and $CSP_1(H)$ are all polynomial or all NP-complete.

Proof. The problem $CSP_1(H)$ is polynomial or NP-complete by Theorem 2.1 because the pairwise-set-full structure H is by definition three-inclusive. Every instance of $CSP^*_{\{1,2\}}(H)$ is an instance of $CSP_{\{1,2\}}(H)$, and every instance of $CSP_{\{1,2\}}(H)$ is an instance of $CSP_1(H)$, so if $CSP_1(H)$ is polynomial, then all three problems are polynomial.

If, on the other hand, $CSP_1(H)$ is NP-complete, then by Corollary 2.3 so is $CSP_1(K)$ where K is obtained from H by deleting all unary relations $S \in \mathcal{U}(H)$ with $|S| > 3$. Since each instance of $CSP_1(K)$ is 1-full, we can prove along the lines of the first part of the proof of Theorem 2.1, that $CSP_1(K)$ is polynomially equivalent to $CSP_1(K')$ where K' is sub-unary. To simplify the notation, let us assume that K itself is sub-unary.

For every pair of variables x, y in an instance G of $CSP_1(K)$, we have x and y constrained by unary relations S^G and T^G respectively, such that $|S^K|, |T^K| \leq 3$. Since H and hence also K is pairwise-set-full, there is a binary relation R^K such that $S^K \times T^K \subseteq R^K$. We may thus add the pair (x, y) to R^G without affecting the existence of a solution. Therefore, for every pair of variables x, y in this modified instance G', there exists a binary relation $R^{G'}$ such that $(x, y) \in R^{G'}$. Hence G' is an instance of $CSP_{\{1,2\}}(K)$, and also of $CSP_{\{1,2\}}(H)$, and thus $CSP_{\{1,2\}}(H)$ is NP-complete.

We can transform the instance G of $CSP_{\{1,2\}}(H)$ into an instance G' of $CSP^*_{\{1,2\}}(H)$, in which each pair (x, y) occurs in only one binary relation R_i, as one of $(x, y) \in R_i$ or $(y, x) \in R_i$. Let r be the number binary of relations in H. For the transformation we shall use an edge coloured complete bipartite graph B with the following properties:

- B has s of vertices in each part;
- each edge of B has one of $2r$ colors;
- between any $s/3$ vertices in one part of B and any $s/3$ vertices in the other part of B, there is an edge of each of the $2r$ colours.

Given B, we replace each variable x in the instance with s variables x_i. Two variables x and y are joined by some $k \leq 2r$ choices of binary relations out of the r binary relations with two possible choices of orientation (from x to y or from y to x) for these binary relations. We may thus join the x_i to the y_i with a bipartite graph obtained from B with edges of k colors, by replacing the remaining $2r - k$ colors with some of the k chosen colors. A variable x is constrained by a unary relation $S \subseteq D^H$ with $|S| \leq 3$, and since H is pairwise-set-full, there exists a binary relation R in H such that $S \times S \subseteq R$. We add the pairs (x_i, x_j) to R, for all the copies x_i and x_j of x. We now have an instance G' of $CSP^*_{\{1,2\}}(H)$. In a solution, out of the s variables x_i, at least $s/3$ of them will have the same value from S, which can then be used as a value for the original variable x. The k binary relations involving x and y appear between the $s/3$ corresponding choices of x_i and the $s/3$ corresponding choices of y_i, so a solution to G' gives a solution to G. This completes the reduction and shows that $CSP^*_{\{1,2\}}(H)$ is NP-complete as well.

It remains to show that such a bipartite graph B exists, for any r. (Note that we may choose s as large as we want.) Given a complete bipartite graph with s vertices on each side, assign a color out of $2r$ colors to each edge uniformly at random. The probability that two sets of size $s/3$ will miss one of the $2r$ colors is at most $2r(1-1/(2r))^{(s/3)^2}$, and the number of choices of such subsets is $\binom{s}{s/3}^2$, so it suffices to choose s large enough so that $2r(1-1/(2r))^{(s/3)^2}\binom{s}{s/3}^2 < 1/2$ to guarantee that with probability at least $1/2$ the complete bipartite graph B will be assigned $2r$ edge colors with the desired property. (The quantity on the left is easily seen to converge to 0 with increasing s, for any constant r.) $\square$

Thus pairwise-set-full constraint satisfaction problems enjoy dichotomy when restricted to either the pairwise full or the strictly pairwise full instances G. We have not succeeded in obtaining a similar dichotomy for three-inclusive constraint satisfaction in general. We can, however, use Theorem 3.1 to obtain a quasi-dichotomy, because of the following result.

Theorem 3.2 *Let H be a three-inclusive structure, and let $\mathcal{H}$ be the set of all pairwise-set-full structures H' with the same domain and the same non-unary relations as H, and such that each $S \in \mathcal{U}(H')$ with $|S| \leq 3$ is also in $\mathcal{U}(H)$.*

If all $CSP_1(H')$ for $H' \in \mathcal{H}$ are polynomial, then both $CSP_{\{1,2\}}(H)$ and $CSP^*_{\{1,2\}}(H)$ are quasi-polynomial. If some $CSP_1(H')$ for $H' \in \mathcal{H}$ is NP-complete, then both $CSP_{\{1,2\}}(H)$ and $CSP^*_{\{1,2\}}(H)$ are NP-complete.

Corollary 3.3 *For each three-inclusive structure H, the problems $CSP_{\{1,2\}}(H)$ and $CSP^*_{\{1,2\}}(H)$ are both NP-complete or both quasi-polynomial.*

Proof. We shall focus on proving the theorem for $CSP_{\{1,2\}}(H)$; the proof for $CSP^*_{\{1,2\}}(H)$ is similar.

We shall transform $CSP_{\{1,2\}}(H)$ to problems $CSP_1(H')$ for $H' \in \mathcal{H}$, in such a way that if any of these problems is NP-complete then so is $CSP_{\{1,2\}}(H)$. In the transformation, each instance G of $CSP_{\{1,2\}}(H)$ with n variables gives rise to $n^{O(\log n)}$ instances G' of problems $CSP_1(H')$, so that if all these problems are polynomial then $CSP_{\{1,2\}}(H)$ is quasi-polynomial.

The problems $CSP_1(H')$ will have a pairwise-set-full structure H' obtained from H by repeatedly replacing certain unary relations S with all proper subsets $T \subset S$. In fact, each structure $H' \in \mathcal{H}'$ will be *strongly pairwise-set-full*, in the sense that it is three-inclusive and for any unary relations S_1, S_2 (not just those with $|S_i| \leq 3$) there is a binary relation R with $S_1 \times S_2$ or $S_2 \times S_1$ contained in R.

To obtain these $n^{O(\log n)}$ instances G' (and the problems $CSP_1(H')$ they correspond to), we proceed as follows. Let A and B be two subsets of D^H such that no binary relation of H contains $A \times B$ or $B \times A$. Since we want to make our structure strongly pairwise-set-full, we shall ensure that there is no $S \in \mathcal{U}(H)$ that contains A, or no $T \in \mathcal{U}(H)$ that contains B. (Later on, we shall keep ensuring this property for other pairs A, B of sets, on structures H' previously obtained from H.)

Let L be the set of variables (elements of D^G) which are in unary constraints S^G such that $A \subseteq S^H$, and let M be the set of variables (elements of D^G) in unary constraints T^G such that $B \subseteq T^H$. Let $\ell = |L|$, $m = |M|$, and let r again denote the number of binary relations in H. For every variable $v \in L$ there exists a binary relation R_i such that (v,w) or (w,v) is in R_i^G for at least $m/(2r)$ variables $w \in M$. We now transform the instance G of $CSP_{\{1,2\}}(H)$ (later on, the instance

G' of the current $CSP_{\{1,2\}}(H')$) into $\ell+1$ new problems. For each binary relation R_i we choose variables $a_i, a'_i \in A$, $b_i, b'_i \in B$ so that $(a_i, b_i) \notin R_i^H, (b'_i, a'_i) \notin R_i^H$. In the first derived problem we change, for each variable $v \in L$, the unary constraint S^G to the unary constraint $S^G - a_i$ or $S^G - a'_i$, depending on whether R_i was chosen above for pairs (v, w) or (w, v). In the remaining ℓ derived problems, we change, for one variable $v \in L$, the constraint S^G to the constraint $\{a_i\}^G$ or $\{a'_i\}^G$, depending on whether R_i was chosen above for pairs (v, w) or (w, v). We also change, for all variables $w \in M$ such that (v, w) (respectively (w, v)) is constrained by R_i^G, the constraint T^G to the constraint $T^G - b_i$ (respectively $T^G - b'_i$). Recall that for each choice of v, the number of such variables w is at least $m/(2r)$. Thus an instance with $(|L|, |M|) = (\ell, m)$ has been replaced with one instance with $(|L|, |M|) = (0, m)$ and ℓ instances with $(|L|, |M|) = (\ell-1, m(1-1/(2r)))$. Repeating this process $-\log m/\log(1-1/(2r)) = O(\log n)$ times, we obtain $n^{O(\log n)}$ instances that each have $\ell = 0$ or $m = 0$, that is, either no unary constraint contains A or no unary constraint contains B. The problem $CSP_{\{1,2\}}(H)$ is correspondingly replaced by problems $CSP_{\{1,2\}}(H')$. Note that no unary constraint in H' contains A or no unary constraint in H' contains B, and all the unary constraints in H' are subsets of unary constraints in H.

We repeat this process for each pair of sets (A, B) as above. In the end, we obtain $n^{O(\log n)}$ instances, and corresponding structures H', such that H' is strongly pairwise-set-full. If all such $CSP_1(H')$ are polynomial, then we can solve each of these $n^{O(\log n)}$ resulting problems, obtaining a quasi-polynomial algorithm for $CSP_{\{1,2\}}(H)$. Otherwise, at least one such $CSP_1(H')$ is NP-complete. Let H'' be obtained from H' by removing all unary constraints $S^{H'}$ with $|S^{H'}| > 3$. By Corollary 2.3 the pairwise-set-full constraint satisfaction problem $CSP_1(H'')$ is also NP-complete; note that all the the unary constraints of H'' are in $\mathcal{U}(H)$. By Theorem 3.1 the problem $CSP_{\{1,2\}}(H'')$ is also NP-complete, and its instances are also instances of the original problem $CSP_{\{1,2\}}(H)$, so this problem is also NP-complete. $\square$

Recall the list matrix partition problems $CSP^*_{1,2}(H)$ defined in the Introduction, and studied in [6, 17, 18, 24, 27]. We obtain the following corollaries. (The first corollary has been anticipated in [17].)

Corollary 3.4 *Each list matrix partition problem is NP-complete or quasi-polynomial.*

We call a list matrix partition problem $CSP^*_{1,2}(H)$ *three-inclusive*, if the instances are graphs G with lists of size at most three.

Corollary 3.5 *Each three-inclusive list matrix partition problem is NP-complete or quasi-polynomial.*

The analogue of Corollary 2.3 follows from Theorems 3.1 and 3.2 by an application of Corollary 2.3.

Corollary 3.6 *If H and K are pairwise-set-full structures on the same domain $D^H = D^K$, which have the same non-unary relations and the same unary relations S with $|S| \leq 3$, then $CSP_{\{1,2\}}(H)$ and $CSP_{\{1,2\}}(K)$ are both NP-complete or both polynomial.*

If H and K are three-inclusive structures on the same domain $D^H = D^K$, which have the same non-unary relations and the same unary relations S with $|S| \leq 3$, then $CSP_{\{1,2\}}(H)$ and $CSP_{\{1,2\}}(K)$ are both NP-complete or both quasi-polynomial.

The same conclusions hold for the strict versions $CSP^*_{\{1,2\}}(H)$ and $CSP^*_{\{1,2\}}(K)$.

4 Restriction to W-Full (and Strictly W-Full) Instances

Theorems 3.1 and 3.2 generalize to W-full structures, and the proofs are similar. Here we briefly state the results, and describe the additional effort required to prove them.

Here is the generalization of Theorem 3.1.

Theorem 4.1 *Let H be a W-set-full structure.*

The problems $CSP_W(H), CSP^*_W(H)$, and $CSP_1(H)$ are all polynomial or all NP-complete.

Proof. We prove the theorem for $W = \{1, w\}$ for a single $w \geq 2$. For general W we can carry out this proof for each value $w \in W$ with $w \geq 2$ in turn. Thus we assume $W = \{1, w\}$ and proceed as in Theorem 3.1. If $CSP_1(H)$ is polynomial, then trivially both $CSP^*_W(H)$ and $CSP_W(H)$ are polynomial. If $CSP_1(H)$ is NP-complete, then by Corollary 2.3 we may restrict $CSP_1(H)$ to three possible values in H per variable (using the unary relations) and still obtain an NP-complete problem. Again, for every collection of distinct variables $x_1, \ldots, x_w$ in D^G, where x_i has at most three possible values given by S_i^H, there exists a relation R^H and a permutation σ such that $S^H_{\sigma(1)} \times \cdots \times S^H_{\sigma(w)} \subseteq R^H$ (since H is W-set-full), so the instance G can be made W-full as before, and thus becomes an instance of the W-set-full problem $CSP_W(H)$. Therefore $CSP_W(H)$ is also NP-complete.

It remains to enforce strictness to obtain an instance of $CSP^*_W(H)$. As in Theorem 3.1, we make s copies of each variable, and use an auxiliary construction to define the relations on this enlarged set of variables. Specifically, we apply Lemma 4.2 below, with $t = w!r$ (corresponding to the $w!$ choices of permutations σ and r choices of relations R^H in the definition of strictly W-full structures). We then obtain an instance equivalent to G that is strictly W-full, whence $CSP^*_W(H)$ is also NP-complete. □

Lemma 4.2 *Let K_{ws} be the structure with ws elements x_{ij}, $1 \leq i \leq w$, $1 \leq j \leq s$ and relation R consisting of all w^s tuples $(x_{1i_1}, \ldots, x_{wi_w})$. For every pair of integers $w \geq 2$ and $t \geq 1$, there exists an integer $s \geq 1$ and a coloring of the tuples in K_{ws} with t colors such that for every choice of w subsets $S_1, \ldots, S_w$ with $S_i \subseteq \{x_{i1}, \ldots, x_{is}\}$ and $|S_i| \geq s/3$, we have that $S_1 \times \cdots \times S_w$ contains tuples of all w-colors.*

Proof. Assign a color out of t colors to each tuple uniformly at random. The probability that a choice of w sets S_i with $|S_i| \geq s/3$ will miss one of the t colors is at most $t(1-1/t)^{(s/3)^t}$, and the number of choices of such subsets is $\binom{s}{s/3}^t$, so it suffices to choose s large enough so that $t(1-1/t)^{(s/3)^w}\binom{s}{s/3}^w < 1/2$ to guarantee that with probability at least $1/2$ the structure K_{ws} will be assigned t tuple colors with the desired property. □

The generalization of Theorem 3.2 takes the following form.

Theorem 4.3 *Let H be a three-inclusive structure, and let $\mathcal{H}$ be the set of all W-set-full structures H' that have the same non-unary relations as H, and such that each $S \in \mathcal{U}(H')$ with $|S| \leq 3$ is also in $\mathcal{U}(H)$.*

If all $CSP_1(H')$ for $H' \in \mathcal{H}$ are polynomial, then both $CSP_W(H)$ and $CSP^*_W(H)$ are quasi-polynomial. If some $CSP_1(H')$ for $H' \in \mathcal{H}$ is NP-complete, then both $CSP_W(H)$ and $CSP^*_W(H)$ are NP-complete.

Corollary 4.4 *For each three-inclusive structure H, the problems $CSP_W(H)$ and $CSP^*_W(H)$ are both NP-complete or both quasi-polynomial.*

In the proof of Theorem 4.3 we shall again treat each $w \in W$ in turn, thus we shall assume that $W = \{1, w\}$.

Let G be a W-full structure, let $L_1, \ldots, L_w$ be (not necessarily disjoint) nonempty subsets of D^G, and let $\alpha = 1/(w!r)$, where r is the number of w-ary relations R_i^G. Say that a rooted tree T is a *select-tree* if the vertices of T at depth $1 \leq i \leq w$ correspond to the variables in L_i, with the root ρ at depth 0. For a w-ary relation R_i^G, a permutation σ of $\{1, \ldots, w\}$, and $0 < \beta < 1$, let a (R_i, σ, β)-*tree* be a select-tree T such that each path $(\rho, x_1, \ldots, x_w)$ in T has $(x_{\sigma(1)}, \ldots, x_{\sigma(w)}) \in R_i^G$ or has $x_i = x_j$ for some $1 \leq i < j \leq w$, and each vertex of T at depth $0 \leq i < w$ has at least $\beta|L_{i+1}|$ children.

Lemma 4.5 *For every W-full structure G and sets $L_1, \ldots, L_w$ as above, there exist R_i and σ such that G has a (R_i, σ, β)-tree for $\beta = \alpha/(4w)$.*

Proof. Let $z = \ell_1 \cdots \ell_w$ for $\ell_i = |L_i|$. There exists a relation R_i^G and a permutation σ of $\{1, \ldots, w\}$ such that at least $\alpha z = z/(w!r)$ tuples $(x_1, \ldots, x_w)$ with $x_i \in L_i$ have $(x_{\sigma(1)}, \ldots, x_{\sigma(w)}) \in R_i^G$ or have $x_i = x_j$ for some $1 \leq i < j \leq w$. By Markov's inequality, there are at least $\beta \ell_1$ elements $x \in L_1$ such that the number of w-tuples $(x_1, \ldots, x_w)$ with $x = x_1$ and $x_i \in L_i$ having $(x_{\sigma(1)}, \ldots, x_{\sigma(w)}) \in R_i^G$, or having $x_i = x_j$ for some $1 \leq i < j \leq w$, is at least $\alpha'(z/\ell_1)$ for $\alpha' = (\alpha - \beta)/(1 - \beta)$, for any β with $0 < \beta < \alpha$. (Markov's inequality states that a nonnegative random variable X has value at most $tE(X)$ with probability at least $1 - \frac{1}{t}$, for each choice of $t > 1$.) By our special choice of β, we have $\alpha''(1 - 1/(2(w-1))) \geq (\alpha'' - \beta)/(1 - \beta)$, for each $\alpha'' \geq \alpha/2$. Thus repeatedly replacing α by α' a total of $w - 1$ times reduces α by a factor at least $(1 - 1/(2(w-1)))^{w-1} \geq 1/2$, so that at any point in time the new value α'' satisfies $\alpha'' \geq \alpha/2$. We therefore obtain at least $\beta \ell_1$ children of the root, and for each chosen child x, we have z replaced with z/ℓ_1 and αz replaced with $\alpha'(z/\ell_1)$, giving again at least $\beta \ell_2$ children of x, and by induction, each vertex of T at depth $0 \leq i < w$ has at least $\beta \ell_{i+1}$ children for $\beta = \alpha/(4w)$. $\square$

Proof of Theorem 4.3. We again focus on proving the statement for $CSP_W(H)$, the proof for $CSP_W^*(H)$ being similar. We proceed as in Theorem 3.2. Assume H is a three-inclusive structure, and transform $CSP_W(H)$ to problems $CSP_1(H')$, for W-set-full structures $H' \in \mathcal{H}$, such that if any of these problems is NP-complete then $CSP_W(H)$ is also NP-complete. In the transformation, each instance G of $CSP_W(H)$ gives rise to $n^{O(\log n)}$ instances G' of the problems $CSP_1(H')$, so that if all these problems are polynomial then $CSP_W(H)$ is quasi-polynomial. Recall that we assume $W = \{1, w\}$. The problems $CSP_1(H')$ will again have a W-set-full structure H' obtained from H by repeatedly replacing certain unary relations S with all proper subsets $T \subset S$.

To obtain these $n^{O(\log n)}$ instances G' (and the problems $CSP_1(H')$ they correspond to), we shall ensure that if some elements $a_{ij\sigma} \in D^H$, one for each of the r w-ary relations R_i^H, $1 \leq j \leq w$ and each fixed permutation σ on $\{1, \ldots, w\}$, are such that $(a_{i\sigma(1)\sigma}, \ldots, a_{i\sigma(w)\sigma}) \notin R_i^H$, and we let $A_j = \{a_{ij\sigma} : 1 \leq i \leq r\}$, then for some $1 \leq j \leq w$ the unary constraints on the instances G' (and corresponding structures H', both described below) do not contain A_j. This will guarantee that we obtain a W-set-full problem.

Let L_j be the set of variables with unary constraints S^G such that $A_j \subseteq S^H$ for $1 \leq j \leq w$, and let $\ell_j = |L_j|$. By Lemma 4.5 G has an (R_i, σ, β) tree T for $\beta = 1/(4w!wr)$. For every solution f to the instance G, there must exist a path $(\rho, x_1, \ldots, x_{t-1})$ in T with $1 \leq t \leq w$, with the tree vertices x_u corresponding to distinct elements of D^G, that gives to x_u value $a_{iu\sigma}$ for $1 \leq u < t$, but does not give value $a_{it\sigma}$ to any child x_t of x_{t-1} corresponding to an element distinct from the elements for these x_u, $1 \leq u < t$.

Since the number of such children of x_{t-1} is at least $\beta \ell_t$, and the solution f belongs to a family of solutions that assigns the values $a_{iu\sigma}$ for $1 \leq u < t$ and removes the value $a_{it\sigma}$ from the lists of these children, we obtain a new instance with ℓ_t replaced by $(1-\beta)\ell_t$. This simplification can be performed at most $-\log(\ell_1\cdots\ell_w)\log(1-\beta) \le -w\log n\log(1-1/(4w!wr)) = O(\log n)$ times, and the number of possible choices of paths $(\rho, x_1, \ldots, x_{t-1})$ for each iteration is at most $2\ell_1\cdots\ell_w \le 2n^w = n^{O(1)}$, so we obtain $n^{O(\log n)}$ instances that have some $\ell_j = 0$, thus satisfying the condition for the corresponding A_j.

We repeat this process for each tuple of sets $(A_1, \ldots, A_w)$ as above. In the end, we obtain $n^{O(\log n)}$ instances (and corresponding structures H') that satisfy the stated condition for each tuple $(A_1, \ldots, A_w)$. Thus for each resulting instance G' of $CSP_1(H')$, the structure H' is strongly W-set-full. If all such $CSP_1(H')$ are polynomial, then we can solve each of these $n^{O(\log n)}$ resulting problems. If one such $CSP_1(H')$ is NP-complete, then by Corollary 2.3 the W-set-full structure H'', obtained from H' by restricting all unary relations to subsets of size at most three, yields a problem $CSP_1(H'')$ that is also NP-complete, and all these subsets of size at most three are in the original structure H. By Theorem 4.1 the problem $CSP_W(H'')$ is NP-complete, and since the instances of $CSP_W(H'')$ are also instances of $CSP_W(H)$, the problem $CSP_W(H)$ is also NP-complete. □

We also obtain the analogue of Corollary 3.6, which follows from Theorems 4.1 and 4.3 by an application of Corollary 2.3.

Corollary 4.6 *If H and K are W-set-full structures on the same domain $D^H = D^K$, which have the same non-unary relations and the same unary relations S with $|S| \le 3$, then $CSP_W(H)$ and $CSP_W(K)$ are both NP-complete or both polynomial.*

If H and K are three-inclusive structures on the same domain $D^H = D^K$, which have the same non-unary relations and the same unary relations S with $|S| \le 3$, then $CSP_W(H)$ and $CSP_W(K)$ are both NP-complete or both quasi-polynomial.

The same conclusions hold for the strict versions $CSP_W^*(H)$ and $CSP_W^*(K)$.

As in the proof of Theorem 4.1, every three-inclusive problem $CSP_W(H)$ can be transformed into a polynomially equivalent problem $CSP_W^*(H)$. However, it is not clear how to accomplish the converse transformation in general; so the latter family of problems may define a more general family, up to polynomial time equivalence. But the two families of problems are the same, up to quasi-polynomial time equivalence, since all problems considered are either quasipolynomial (hence equivalent up to quasi-polynomial reductions) or NP-complete (hence equivalent up to polynomial reductions).

5 Bounded Degree Problems

Here we consider restricting the three-inclusive constraint satisfaction problems $CSP_1(H)$ to instances G with bounded degree. We are lead to propose the following conjecture.

Conjecture 5.1 *Every three-inclusive list constraint satisfaction problem $CSP_1(H)$ that is NP-complete, remains NP-complete when restricted to instances of degree three.*

We denote by $CSP_1^3(H)$ the restriction of $CSP_1(H)$ to instances of degree three. We now briefly explain four special cases in which the conjecture holds. All are based on existing results.

We shall say that a three-inclusive constraint satisfaction problem $CSP_1(H)$ *d-simulates* a three-inclusive constraint satisfaction problem $CSP_1(H')$, if for every relation $R_i^{H'}$, there is an instance G_i of $CSP_1(H)$ of degree at most d containing ℓ_i particular variables $x_1, \ldots, x_{\ell_i}$ of degree at most one, such that $(a_1, \ldots, a_{\ell_i}) \in R_i^{H'}$ if and only if there exists a homomorphism f of G_i to H such that each $f(x_i) = a_i$. If $CSP_1(H)$ d-simulates $CSP_1(H')$ with $d \leq 3$, then there is a polynomial reduction of the problem $CSP_1^3(H')$ to $CSP_1^3(H)$, obtained by substituting each occurrence of a relation $R_i^{H'}$ in an instance G' of $CSP_1^3(H')$ by a copy of the instance G_i corresponding to R_i, thus obtaining an instance G of $CSP_1^3(H)$. Therefore showing the conjecture for $CSP_1^3(H')$ implies the conjecture for $CSP_1^3(H)$.

In the first case, we focus on the case of *Boolean constraint satisfaction problems*, i.e., $CSP_1(H)$, where D^H has two elements $(0, 1)$. The theorem stated below was first proved by Dalmau and Ford [7], and by Feder and Ford [11], using a result of Feder [9]. We briefly explain the proof as it is the basis for the other cases.

Theorem 5.2 *[7, 11] Conjecture 5.1 holds for Boolean constraint satisfaction problems.*

Proof. In other words, we claim that if for a three-inclusive H with $D = D^H = \{0, 1\}$ the problem $CSP_1(H)$ is NP-complete, then so is its restriction $CSP_1^3(H)$ to instances of degree three. It is shown in [9] that each NP-complete case of Boolean constraint satisfaction problems $CSP_1(H)$ 2-simulates the problem $CSP_1(H')$ which contains the ternary relation $D^3 - \{(0, 0, 0)\}$ (corresponding to the disjunction $x \vee y \vee z$), the ternary relation $D^3 - \{(1, 1, 1)\}$ (corresponding to the disjunction $\overline{x} \vee \overline{y} \vee \overline{z}$), and the binary relation $D^2 - \{(1, 0)\}$ (corresponding to the implication $x \rightarrow y$).

This simulation yields a reduction from the NP-complete problem of 3-satisfiability to the problem $CSP_1^3(H')$, as we may include a variable x in any number k of clauses by using k auxiliary variables $x_1, x_2, \ldots, x_k$ with implications $x_1 \rightarrow x_2, x_2 \rightarrow x_3, \ldots, x_{k-1} \rightarrow x_k, x_k \rightarrow x_1$. These implications involve two occurrences for each x_i, and the third occurrence of each of these k variables x_i that all take the same value can be used for the k occurrences of x in clauses $x \vee y \vee z$ or $\overline{x} \vee \overline{y} \vee \overline{z}$. $\square$

In the second case, we focus on three-inclusive constraint satisfaction problems $CSP_1(H)$ in which all $S \in \mathcal{U}(H)$ have $|S| \leq 2$. These problems correspond to list constraint satisfaction problems with lists of size at most two, and we call these problems *small-list constraint satisfaction problems.*

Theorem 5.3 *Conjecture 5.1 holds for small-list constraint satisfaction problems.*

Proof. It is shown in [10] that each NP-complete case of the small-list constraint satisfaction problems 2-simulates a problem $CSP_1(H')$ for a structure H' containing the unary relations $\{0_i, 1_i\}$, for $1 \leq i \leq r$, and containing the binary relations $(\{0_i, 1_i\} \times \{0_j, 1_j\}) - \{(1_i, 0_j)\}$ (corresponding to the implications $x_i \rightarrow x_j$) or the binary relations $(\{0_i, 1_i\} \times \{0_j, 1_j\}) - \{(1_i, 0_j), (0_i, 1_j)\}$ (corresponding to the equalities $x_i = x_j$) for each $1 \leq i, j \leq r$, $i \neq j$. It is also shown the [10] that the problem $CSP_1(H)$ with $D^H = \{0, 1\}$, obtained from $CSP_1(H')$ by identifying all sets $\{0_i, 1_i\}$ with $\{0, 1\}$, is NP-complete.

We may thus represent kr occurrences of a Boolean variable x, with k of these occurrences having unary relation $\{0_i, 1_i\}$ for each $1 \leq i \leq r$, by a cycle of implications $x_i \rightarrow x_j$ or equalities $x_i = x_j$, as in the proof of Theorem 5.2, involving variables x_{ij} having unary relation $\{0_i, 1_i\}$, for $1 \leq i \leq r$ and $1 \leq j \leq k$. This completes the reduction from the NP-complete problem $CSP_1(H)$ obtained by identifying all $\{0_i, 1_i\}$ with $\{0, 1\}$ to $CSP_1^3(H')$. $\square$

The third case concerns list homomorphism problems for graphs (with loops allowed). The problem $CSP_1(H)$, where H is three-inclusive and consists of a single symmetric binary relation, will be called a *graph list homomorphism problem.* The following theorem has been proved in [15].

Theorem 5.4 *[15] If H is a bi-arc graph, then $CSP_1^3(H)$ is polynomial; otherwise, $CSP_1^3(H)$ is NP-complete.*

Since this classification agrees exactly with that for $CSP_1(H)$ from [14], we obtain the desired corollary.

Corollary 5.5 *Conjecture 5.1 holds for graph list homomorphism problems.*

The last case deals with N-free binary relations. A binary relation B is called *N-free* if $(x,z),(y,z),(y,t) \in B$ imply either $x = y$ or $z = t$. A *binary N-free constraint satisfaction problem* is a three-inclusive problem $CSP_1(H)$ where $\mathcal{U}(H)$ contains all S with $|S| \leq 3$, all binary relations of H are N-free, and H has no relations of higher arity.

Theorem 5.6 *Conjecture 5.1 holds for binary N-free constraint satisfaction problems.*

Proof. It is shown in [19] that each NP-complete binary N-free constraint satisfaction problem 3-simulates a problem $CSP_1(H)$ containing the relation $\{(1_A,0_B,0_C),(0_A,1_B,0_C),(0_A,0_B,1_C)\}$, with $|\{0_A,1_A\}| = |\{0_B,1_B\}| = |\{0_C,1_C\}| = 2$.

This relation may further be used to 3-simulate a problem $CSP_1(H')$ that also contains the relation $\{(1_A,0_B),(0_A,1_B)\}$ (corresponding to the inequality $x_A \neq x_B$), by setting x_C to value 0_C, and similarly $\{(1_B,0_C),(0_B,1_C)\}$ (corresponding to the inequality $x_B \neq x_C$); and also contains the relation $\{(0_A,0_C),(1_A,1_C)\}$ (corresponding to the equality $x_A = x_C$), obtained by combining $x_A \neq x_B$ with $x_B \neq x_C$, and similarly the relation $\{(0_A,0_B),(1_A,1_B)\}$ (corresponding to the equality $x_A = x_B$). Thus we may set $x_{A1} = \cdots = x_{Ak} = x_{B1} \cdots x_{Bk} = x_{C1} \cdots x_{Ck}$ as a chain of equalities that uses only two of the three allowed occurrences of each variable. This allows us to identify the three unary relations $\{0_i,1_i\}$ with a single unary relation $\{0,1\}$, with k occurrences per variable, for the NP-complete Boolean constraint satisfaction problem one-in-three SAT, with relation $\{(1,0,0),(0,1,0),(0,0,1)\}$ (obtained from $\{(1_A,0_B,0_C),(0_A,1_B,0_C),(0_A,0_B,1_C)\}$ by the identification $x_A = x_B = x_C$). Thus $CSP_1^3(H')$ and $CSP_1^3(H)$, with degree at most three, are also NP-complete. $\square$

We are grateful to the referees for guiding us to present these results more clearly.

References

[1] J. Bang-Jensen, P. Hell and G. MacGillivray, "On the complexity of colouring by superdigraphs of bipartite graphs," Discrete Math. 109 (1992) 27–44.

[2] A.A. Bulatov, "Tractable conservative constraint satisfaction problems," In Proceedings of the 18th IEEE Annual Symposium on Logic in Computer Science (LICS 2003), 321–330.

[3] A.A. Bulatov, "A dichotomy constraint on a three-element set," In Proceedings of the 43rd IEEE Symposium on Theory of Computing (2002), 649–658.

[4] A.A. Bulatov and P. Jeavons, "Algebraic structures in combinatorial problems," Technical Report MATH-AL-4-2001m Technishe universität Dresden, Germany, 2001.

[5] A. Bulatov, A. Krokhin, and P. Jeavons, "Classifying complexity of constraints using finite algebras," SIAM J. on Computing 34 (2005) 720–742.

[6] K. Cameron, E. E. Eschen, C. T. Hoang and R. Sritharan, "The list partition problem for graphs," Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) 2004, 391–399.

[7] V. Dalmau and D. Ford, "Generalized satisfiability with k occurrences per variable: a study through delta-matroid parity," Mathematical Foundations of Computer Science (MFCS 2003), Lecture Notes in Computer Science 2747, Springer 2003, 358–367.

[8] R. Dechter, "Constraint networks," In Encyclopedia of Artificial Intelligence, 1992, 276–285.

[9] T. Feder, "Fanout limitations on constraint systems," Theoretical Computer Science 255 (2001) 281–293.

[10] T. Feder, "Homomorphisms to oriented cycles and k-partite satisfiability," SIAM J. Discrete Math. 14 (2001) 471–480.

[11] T. Feder and D. Ford, "Classification of bipartite Boolean constraint satisfaction through delta-matroid intersection," manuscript.

[12] T. Feder and P. Hell, "List homomorphisms to reflexive graphs," J. Comb. Theory Series B 72 (1998) 236–250.

[13] T. Feder, P. Hell, and J. Huang, "List homomorphisms and circular arc graphs," Combinatorica 19 (1999) 487–505.

[14] T. Feder, P. Hell, and J. Huang, "Bi-arc graphs and the complexity of list homomorphisms," J. Graph Theory 42 (1999) 61–80.

[15] T. Feder, P. Hell, and J. Huang, "List homomorphisms of graphs with bounded degrees," to appear in Discrete Math.

[16] T. Feder, P. Hell, and J. Huang, "List homomorphisms to reflexive digraphs," manuscript.

[17] T. Feder, P. Hell, S. Klein, and R. Motwani, "Complexity of list partitions," SIAM J. Discrete Mathematics 16 (2003) 449-478.

[18] T. Feder, P. Hell, D. Kral, and J. Sgall, "Two algorithms for list matrix partitions." Proc. 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) 2005, 870–876.

[19] T. Feder, F. Madelaine, and I.A. Stewart, "Dichotomies for classes of homomorphism problems involving unary functions," Theoretical Computer Science 314 (2004), 1–43.

[20] T. Feder and M.Y. Vardi, "The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory," SIAM J. Comput. 28 (1998) 236–250.

[21] A. Galluccio, P. Hell, J. Nešetřil, "The complexity of H-colouring of bounded degree graphs", Discrete Math. 222 (2000) 101 - 109.

[22] M. C. Golumbic, **Algorithmic Graph Theory and Perfect Graphs**, Academic Press, New York, 1980.

[23] P. Hell, "From graph colouring to constraint satisfaction - there and back again", to appear.

[24] P. Hell, S. Klein, L. Tito Nogueira, and F. Protti, “Partitioning chordal graphs into independent sets and cliques”, Discrete Applied Math. 141 (2004) 185-194.

[25] P. Hell and J. Nešetřil, “Counting list homomorphisms and graphs with bounded degrees”, in **Graphs, Morphisms and Statistical Physics** (J. Nešetřil and P. Winkler, eds.) DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Volume 63 (2004) 105 – 112.

[26] P. Hell and J. Nešetřil, “On the complexity of H-coloring,” J. Comb. Theory, Series B 48 (1990), 92–110.

[27] P. Hell and J. Nešetřil, **Graphs and Homomorphisms**, Oxford University Press 2004.

[28] P. Jeavons, “On the algebraic structure of combinatorial problems,” Theoretical Computer Science 200 (1998) 185–204.

[29] P. Jeavons, D. Cohen, and M. Gyssens, “Closure properties of constraints,” Journal of the ACM 44 (1997) 527–548.

[30] R. E. Ladner, “On the structure of polynomial time reducibility”, J. Assoc. Comput. Mach. 22 (1975) 155 – 171.

[31] P. Lincoln and J. C. Mitchell, “Algorithmic aspects of type inference with subtypes,” in Conf. Rec. 19th ACM Symp. on Principles of Programming Languages (1992), 293–304.

[32] V. Kumar, “Algorithms for constraint-satisfaction problems,” AI Magazine 13 (1992), 32–44.

[33] P. Meseguer, “Constraint satisfaction problem: an overview,” AICOM 2 (1989), 3–16.

[34] J. C. Mitchell, “Coercion and type inference (summary),” in Conf. Rec. 11th ACM Symp. on Principles of Programming Languages (1984), 175–185.

[35] N. Megiddo and U. Vishkin, “On finding a minimum dominating set in a tournament”, Theor. Comput. Sci. 61 (1988) 307–316.

[36] U. Montanari, “Networks of constraints: Fundamental properties and applications to picture processing,” Information Sciences 7 (1974) 95–132.

[37] C.H. Papadimitriou and M. Yannakakis, “On limited nondeterminism and the complexity of the V-C dimension”, J. Comput. Sys. Sci. 53 (1996) 161–170.

[38] T. J. Schaefer, “The complexity of satisfiability problems,” Proc. 10th ACM Symp. on Theory of Computing (1978), 216–226.

[39] M.Y. Vardi, “Constraint satisfaction and database theory: a tutorial,” Proceedings of the 19th Symposium on Principles of Database Systems (PODS 2000), 76–85.

[40] M. Wand and P. M. O’Keefe, “On the complexity of type inference with coercion,” in Conf. on Functional Programming Languages and Computer Architecture (1989), 293–298.