

Polynomial-Time Data Reduction for DOMINATING SET*

Jochen Alber[†] Michael R. Fellows[‡] Rolf Niedermeier[§]

Abstract

Dealing with the NP-complete DOMINATING SET problem on graphs, we demonstrate the power of data reduction by preprocessing from a theoretical as well as a practical side. In particular, we prove that DOMINATING SET restricted to planar graphs has a so-called problem kernel of linear size, achieved by two simple and easy to implement reduction rules. Moreover, having implemented our reduction rules, first experiments indicate the impressive practical potential of these rules. Thus, this work seems to open up a new and prospective way how to cope with one of the most important problems in graph theory and combinatorial optimization.

1 Introduction

Motivation. A core tool for practically solving NP-hard problems is data reduction through preprocessing. Weihe [41, 42] gave a striking example when dealing with the NP-complete RED/BLUE DOMINATING SET problem appearing in the context of the European railroad network. In a preprocessing phase, he applied two simple data reduction rules again and again until

*An extended abstract of this work entitled “Efficient Data Reduction for DOMINATING SET: A Linear Problem Kernel for the Planar Case” appeared in the *Proceedings of the 8th Scandinavian Workshop on Algorithm Theory (SWAT 2002)*, Lecture Notes in Computer Science (LNCS) 2368, pages 150–159, Springer-Verlag 2002.

[†]Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13, D-72076 Tübingen, Fed. Rep. of Germany. Email: alber@informatik.uni-tuebingen.de. Work supported by the Deutsche Forschungsgemeinschaft (DFG), research project PEAL (parameterized complexity and exact algorithms), NI 369/1.

[‡]Department of Computer Science and Software Engineering, University of Newcastle, University Drive, Callaghan 2308, Australia. Email: mfellows@cs.newcastle.edu.au.

[§]Contact author. Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13, D-72076 Tübingen, Fed. Rep. of Germany. Email: niedermr@informatik.uni-tuebingen.de. Work partially supported by the Deutsche Forschungsgemeinschaft (DFG), junior research group PIAF (fixed-parameter algorithms), NI 369/4.

no further application was possible. The impressive result of his empirical study was that each of his real-world instances was broken into very small pieces such that for each of these a simple brute-force approach was sufficient to solve the computationally hard problems efficiently and optimally. In this work, we present a new and stronger scenario for data reduction through preprocessing, namely for the NP-complete DOMINATING SET problem, a core problem in combinatorial optimization and graph theory. According to a 1998 survey [28, Chapter 12], more than 200 research papers and more than 30 PhD theses investigate the algorithmic complexity of domination and related problems [39]. Moreover, domination problems occur in numerous practical settings, ranging from strategic decisions such as locating radar stations or emergency services through computational biology to voting systems (see [28, 29, 37] for a survey). Two recent examples for applications of domination problems can be found in [30] (“power domination” in electric networks) and in [40] (“connected domination” in wireless adhoc networks). By way of contrast to the aforementioned example given by Weihe, however, our preprocessing is, on the one hand, more involved to develop, and, on the other hand, it does not only prove its strength through experimentation but, in first place, by theoretically sound means. Thus, we come up with a practically promising as well as theoretically appealing result for computing the domination number of a graph, one of the so far few positive news for this important problem. To some extent our results also complement a recent experimental analysis of heuristic algorithms for DOMINATING SET [38].

Problem definition and status. A k -dominating set D of an undirected graph G is a set of k vertices of G such that each of the rest of the vertices has at least one neighbor in D . The minimum k such that G has a k -dominating set is called the *domination number* of G , denoted by $\gamma(G)$. The DOMINATING SET problem is to decide, given a graph $G = (V, E)$ and a positive integer k , whether $\gamma(G) \leq k$. Due to its NP-completeness and its practical importance, DOMINATING SET has been subject to intensive studies that were concerned with coping strategies to attack its intractability. Among these coping strategies, we find approximation algorithms and (exact) fixed-parameter algorithms. As to approximation results, it is known that DOMINATING SET is polynomial-time approximable with factor $1 + \log |V|$ since the problem is a special case of the MINIMUM SET COVER problem [31]. On the negative side, however, it is known not to be approximable within $(1 - \epsilon) \ln |V|$ for any $\epsilon > 0$ unless $\text{NP} \subseteq \text{DTIME}(n^{\log \log n})$ [21]. When restricted to planar graphs, where it still remains NP-complete [27], however,

a polynomial time approximation scheme (PTAS) is stated [11].¹ There are numerous approximation results for further special instances of DOMINATING SET, cf. [10]. As to fixed-parameter results, the central question is whether the problem is optimally solvable in $f(k) \cdot n^{O(1)}$ time, where $f(k)$ may be an exponentially fast (or worse) growing function in the *parameter* k only and n is the number of graph vertices. Unfortunately, also here the situation seems hopeless—the problem is known to be W[2]-complete [18, 19] which implies fixed-parameter intractability unless very unlikely collapses of parameterized complexity classes occur (see [19] for details). Again, restricting the problem to planar graphs improves the situation. Then, DOMINATING SET is known to be solvable in $O(c^{\sqrt{k}} \cdot n)$ time for $c \leq 4^{6\sqrt{34}}$ [3]² and, alternatively, solvable in $O(8^k \cdot n)$ time [4]. Recently, the upper bound on the constant c was improved to 2^{27} [32] and further to $2^{15.13}$ [25]. As to fixed-parameter complexity, it was open whether DOMINATING SET on planar graphs possesses a so-called problem kernel of linear size, a question we answer affirmatively here.

Results. We provide positive news on the algorithmic tractability of DOMINATING SET through preprocessing. The heart of our results are two relatively simple and easy to implement “reduction rules” for DOMINATING SET. These rules are based on considering local structures within the graph. They produce a reduced graph such that the original graph has a dominating set of size at most k iff the reduced graph has a dominating set of size at most k' for some $k' \leq k$. The point here is that the reduced graph, as a rule, is much smaller than the original graph and, thus, k' is significantly smaller than k because the reduction process usually determines several vertices that are part of an optimal dominating set. In this way, these two reduction rules provide an efficient data reduction through polynomial-time preprocessing. In the case of planar graphs, we actually can prove that the reduced graph consists of at most $335k$ vertices (which is completely independent of the size of the original graph). In fixed-parameter complexity terms, this means that DOMINATING SET on planar graphs possesses a linear size problem kernel. Note, however, that our main concern in analyzing the multiplicative constant 335 was conceptual simplicity for which we deliberately sacrificed the aim to further lower it by way of refined analysis (without changing

¹In [11], only the conceptually much simpler INDEPENDENT SET problem is described in detail.

²Note that in the SWAT 2000 conference version of [3], an exponential base $c = 3^{6\sqrt{34}}$ is stated, caused by a misinterpretation of previous results. The correct worst-case upper bound reads $c = 4^{6\sqrt{34}}$.

the reduction rules). Finally, experimental studies underpin the big potential of the presented reduction rules, leading to graph size reductions of more than 90 percent when experimenting with random planar graphs and so-called Internet graphs [2]. Hence, we conjecture that future algorithms for DOMINATING SET, whether approximation, fixed-parameter, or purely heuristic, should employ data reduction by preprocessing. The point here is that a problem kernel as achieved by our data reduction rules can be the starting point for any algorithmic strategy to apply. This observation is further substantiated by the fact that data reduction by preprocessing plays an eminently important role when hard combinatorial problems are solved in practice.

Relation to previous work. Our data reduction still allows to solve the problem exactly, not only approximately. It is, thus, always possible to incorporate our reduction rules in any kind of approximation algorithm for DOMINATING SET without deteriorating its approximation factor. In this sense, Baker’s PTAS result³ for DOMINATING SET on planar graphs [11] probably has less applicability than the result presented here. This is due to the fact that her scenario including dynamic programming (which we also used when applying our related approach based on tree decompositions [3]) seems to require much computational overhead (including high constant factors in the running time). Our data reduction algorithm is conceptually much simpler and, as a preprocessing method, seems to combine with *any* kind of algorithm working afterwards on the then reduced graph.

Concerning the parameterized complexity of DOMINATING SET on planar graphs, we have the following consequences of our result. First, on the structural side, combining our linear problem kernel with the graph separator approach presented in [6] immediately results in an $O(c^{\sqrt{k}} \cdot k + n^{O(1)})$ DOMINATING SET algorithm on planar graphs (for some constant c). Also, the linear problem kernel directly proves the so-called “Layerwise Separation Property” [5] for DOMINATING SET on planar graphs, again implying an $O(c^{\sqrt{k}} \cdot k + n^{O(1)})$ algorithm. Second, the linear problem kernel improves the time $O(8^k \cdot n)$ search tree algorithm from [4] to an $O(8^k k + n^{O(1)})$ algorithm. We are aware of only one further result that provides a *provable* data reduction by preprocessing in our sense, namely the Nemhauser-Trotter theorem

³There is an ongoing discussion and investigation of the practical usefulness of (most) PTAS results [17, 23]. The problem with PTAS algorithms often is that they require high-degree polynomial running time in order to achieve a reasonably good degree of approximation. Actually, the third author, attending a DIMACS workshop on approximation algorithms held in Princeton in February 2000, remembers one of the speakers asking for any examples where a PTAS really has been applied in practice.

for VERTEX COVER [35, 12, 33]. Their polynomial-time preprocessing employs a maximum matching algorithm for bipartite graphs and provides a reduced graph where at least half of the vertices have to be part of an optimal vertex cover set (also see [13] for details and its implication of a size $2k$ problem kernel). Note, however, that from an algorithmic and combinatorial point of view, VERTEX COVER seems to be a much less elusive problem⁴ than DOMINATING SET is.

Structure of the paper. We start with our two reduction rules based on the neighborhood structure of a single vertex and a pair of vertices, respectively. Here, we also analyze the worst-case time complexity of these reduction rules for planar as well as for general graphs. Afterwards, in the technically most demanding part, we prove that for planar graphs our reduction rules *always* deliver a reduced graph of size $O(\gamma(G))$. Finally, we discuss some experimental findings and give some conclusions and challenges for future work.

2 The Reduction Rules

We present two reduction rules for DOMINATING SET. Both reduction rules are based on the same principle: We explore local structures of the graph and try to replace them by simpler structures. For the first reduction rule, the local structure will be the neighborhood of a single vertex. For the second reduction rule, we will deal with the union of the neighborhoods of a pair of vertices.

2.1 The Neighborhood of a Single Vertex

Consider a vertex $v \in V$ of the given graph $G = (V, E)$. Here and in the following for $v \in V$ let $N(v) := \{u : \{u, v\} \in E\}$ be the *neighborhood* of v . We partition the vertices of $N(v)$ of v into three different sets $N_1(v)$, $N_2(v)$, and $N_3(v)$ depending on what neighborhood structure these vertices have. More precisely, setting $N[v] := N(v) \cup \{v\}$, we define

$$\begin{aligned} N_1(v) &:= \{u \in N(v) : N(u) \setminus N[v] \neq \emptyset\},^5 \\ N_2(v) &:= \{u \in N(v) \setminus N_1(v) : N(u) \cap N_1(v) \neq \emptyset\}, \\ N_3(v) &:= N(v) \setminus (N_1(v) \cup N_2(v)). \end{aligned}$$

⁴For instance, VERTEX COVER has a simple factor 2 approximation algorithm and it has fixed-parameter algorithms of $O(1.29^k + kn)$ running time on general graphs [13, 36].

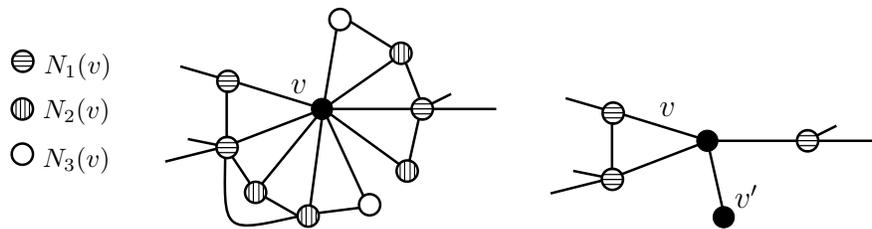


Figure 1: The left-hand side shows the partitioning of the neighborhood of a single vertex v . The right-hand side shows the result of applying Rule 1 to this particular (sub)graph.

An example which illustrates the partitioning of $N(v)$ into the subsets $N_1(v)$, $N_2(v)$, and $N_3(v)$ can be seen in the left-hand diagram of Fig. 1.

Note that, by definition of the three subsets, the vertices in $N_3(v)$ cannot be dominated by vertices from $N_1(v)$. A good candidate for dominating $N_3(v)$ is given by the choice of v . Observing that this indeed is always an optimal choice lies the base for our first reduction rule.

Rule 1. *If $N_3(v) \neq \emptyset$ for some vertex v , then*

- *remove $N_2(v)$ and $N_3(v)$ from G and*
- *add a new vertex v' with the edge $\{v, v'\}$ to G .*

We use the vertex v' as a “gadget vertex” that enforces us to take v (or v') into an optimal dominating set in the reduced graph.

Example 1. *Fig. 1 shows the neighborhood of a vertex v before and after applying Rule 1 to it.*

Lemma 1. *Let $G = (V, E)$ be a graph and let $G' = (V', E')$ be the resulting graph after having applied Rule 1 to G . Then $\gamma(G) = \gamma(G')$.*

Proof. Consider a vertex $v \in V$ such that $N_3(v) \neq \emptyset$. The vertices in $N_3(v)$ can only be dominated by either v or by vertices in $N_2(v) \cup N_3(v)$. But, clearly, $N(w) \subseteq N(v)$ for every $w \in N_2(v) \cup N_3(v)$. This shows that an optimal way to dominate $N_3(v)$ is given by taking v into the dominating set. This is simulated by the “gadget vertex” v' in G' which enforces us to take v (or v') into an optimal dominating set. It is safe to remove $N_2(v) \cup$

⁵For two sets X, Y , where Y is not necessarily a subset of X , we use the convention that $X \setminus Y := \{x \in X : x \notin Y\}$.

$N_3(v)$ since $N(N_2(v) \cup N_3(v)) \subseteq N(v)$, i.e., since the vertices that could be dominated by vertices from $N_2(v) \cup N_3(v)$ are already dominated by v . Hence, $\gamma(G') = \gamma(G)$. \square

Lemma 2. *Rule 1 can be carried out in $O(n)$ time for planar graphs and in $O(n^3)$ time for general graphs.*

Proof. We first discuss the planar case. To carry out Rule 1, for each vertex v of the given planar graph G we have to determine the neighbor sets $N_1(v)$, $N_2(v)$, and $N_3(v)$. By definition of these sets, one easily observes that it is sufficient to consider the subgraph G that is induced by all vertices that are connected to v by a path of length at most two. To do so, we employ a “partial” depth-first search tree of depth two, rooted at v . More precisely, this means that we explore all vertices at distance one from v (i.e., connected to v by an edge in G) and some vertices at distance two from G (to be described in more detail in the following). We perform two phases.

In phase 1, constructing the search tree we determine the vertices from $N_1(v)$. Each vertex of the first level (i.e., distance one from the root v) of the search tree that has a neighbor at the second level of the search tree belongs to $N_1(v)$. Observe that it is enough to stop the expansion of a vertex from the first level as soon as its *first* neighbor in the second level is encountered. Hence, denoting the degree of v by $\deg(v)$, phase 1 takes time $O(\deg(v))$ because there clearly are at most $2 \cdot \deg(v)$ tree edges and at most $O(\deg(v))$ non-tree edges to be explored. The latter holds true since these non-tree edges all belong to the subgraph of G induced by $N[v]$. Since this graph is clearly planar and $|N[v]| = \deg(v) + 1$, the claim follows.

In phase 2, it remains to determine the sets $N_2(v)$ and $N_3(v)$. To get $N_2(v)$, one basically has to go through all vertices from the first level of the above search tree that are not already marked as being in $N_1(v)$ but have at least one neighbor in $N_1(v)$. All this can be done within the planar graph induced by $N[v]$, using the already marked $N_1(v)$ -vertices, in time $O(\deg(v))$. Finally, $N_3(v)$ simply consists of vertices from the first level that are neither marked being in $N_1(v)$ nor marked being in $N_2(v)$. In summary, this shows that for a vertex v the sets $N_1(v)$, $N_2(v)$, and $N_3(v)$ can be constructed in time $O(\deg(v))$.

Once having determined these three sets, the sizes of which all are bounded by $\deg(v)$, it is clear that the possible removal of vertices from $N_2(v)$ and $N_3(v)$ and the addition of a vertex and an edge as required by Rule 1 all can be done in time $O(\deg(v))$. Finally, it remains to analyze the overall complexity of this procedure when going through all n vertices of $G = (V, E)$.

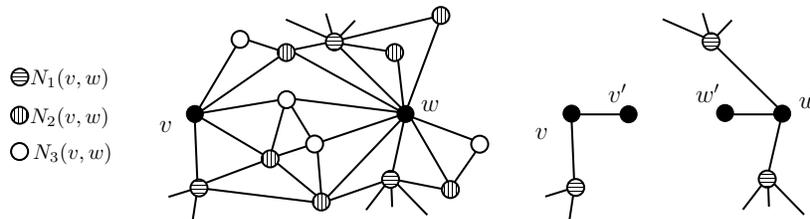


Figure 2: The left-hand side shows the partitioning of a neighborhood $N(v, w)$ of two vertices v and w . The right-hand side shows the result of applying Rule 2, Case 2 to this particular (sub)graph.

But this is easy. The running time can be bounded by $\sum_{v \in V} O(\deg(v))$. Since G is planar, this sum is bounded by $O(n)$, i.e., the whole reduction takes linear time.

For general graphs, the method described above leads to a worst-case cubic time implementation of Rule 1. Here, one ends up with the sum

$$\sum_{v \in V} O((\deg(v))^2) = O(n^3).$$

Note that the size of the graph that is induced by the neighborhood $N[v]$ again is relevant for the time needed to determine the sets $N_1(v)$, $N_2(v)$, and $N_3(v)$. For general graphs, this neighborhood may contain $O((\deg(v))^2)$ many edges. \square

2.2 The Neighborhood of a Pair of Vertices

Similar to Rule 1, we explore the *neighborhood* set $N(v, w) := N(v) \cup N(w) \setminus \{v, w\}$ of two vertices $v, w \in V$. Analogously, we now partition $N(v, w)$ into three disjoint subsets $N_1(v, w)$, $N_2(v, w)$, and $N_3(v, w)$. Setting $N[v, w] := N[v] \cup N[w]$, we define

$$\begin{aligned} N_1(v, w) &:= \{u \in N(v, w) : N(u) \setminus N[v, w] \neq \emptyset\}, \\ N_2(v, w) &:= \{u \in N(v, w) \setminus N_1(v, w) : N(u) \cap N_1(v, w) \neq \emptyset\}, \\ N_3(v, w) &:= N(v, w) \setminus (N_1(v, w) \cup N_2(v, w)). \end{aligned}$$

The left-hand diagram of Fig. 2 shows an example which illustrates the partitioning of $N(v, w)$ into the subsets $N_1(v, w)$, $N_2(v, w)$, and $N_3(v, w)$. Our second reduction rule—compared to Rule 1—is slightly more complicated.

Rule 2. Consider $v, w \in V$ ($v \neq w$) and suppose that $|N_3(v, w)| > 1$. Suppose that $N_3(v, w)$ cannot be dominated by a single vertex from $N_2(v, w) \cup N_3(v, w)$.

Case 1 If $N_3(v, w)$ can be dominated by a single vertex from $\{v, w\}$:

- (1.1) If $N_3(v, w) \subseteq N(v)$ as well as $N_3(v, w) \subseteq N(w)$:
 - remove $N_3(v, w)$ and $N_2(v, w) \cap N(v) \cap N(w)$ from G and
 - add two new vertices z, z' and edges $\{v, z\}$, $\{w, z\}$, $\{v, z'\}$, $\{w, z'\}$ to G .
- (1.2) If $N_3(v, w) \subseteq N(v)$, but not $N_3(v, w) \subseteq N(w)$:
 - remove $N_3(v, w)$ and $N_2(v, w) \cap N(v)$ from G and
 - add a new vertex v' and the edge $\{v, v'\}$ to G .
- (1.3) If $N_3(v, w) \subseteq N(w)$, but not $N_3(v, w) \subseteq N(v)$:
 - remove $N_3(v, w)$ and $N_2(v, w) \cap N(w)$ from G and
 - add a new vertex w' and the edge $\{w, w'\}$ to G .

Case 2 If $N_3(v, w)$ cannot be dominated by a single vertex from $\{v, w\}$:

- remove $N_3(v, w)$ and $N_2(v, w)$ from G and
- add two new vertices v', w' and edges $\{v, v'\}$, $\{w, w'\}$ to G .

Clearly, Cases (1.2) and (1.3) are symmetric to each other. Again, the newly added vertices v' and w' of degree one act as gadgets that enforce us to take v or w into an optimal dominating set. A special situation is given in Case (1.1). Here, the gadget added to the graph G simulates that at least one of the vertices v or w has to be taken into an optimal dominating set.

Example 2. Fig 2 shows an application of Rule 2, Case 2.

Lemma 3. Let $G = (V, E)$ be a graph and let $G' = (V', E')$ be the resulting graph after having applied Rule 2 to G . Then $\gamma(G) = \gamma(G')$.

Proof. Similar to the proof of Lemma 1, we observe that vertices from $N_3(v, w)$ can only be dominated by vertices from $M := \{v, w\} \cup N_2(v, w) \cup N_3(v, w)$. All cases in Rule 2 are based on the fact that $N_3(v, w)$ needs to be dominated. All cases only apply if there is not a *single* vertex in $N_2(v, w) \cup N_3(v, w)$ which dominates $N_3(v, w)$.

We first of all discuss the correctness of Case (1.2) (and similarly obtain the correctness of the symmetric Case (1.3)): If v dominates $N_3(v, w)$ (and w does not) then it is optimal to take v into the dominating set—and at

the same time still leave the option of taking vertex w —than to take any combination of two vertices x, y from the set $M \setminus \{v\}$. It may be that we still have to take w to get a minimum dominating set, but in any case v and w dominate at least as many vertices as x and y . The “gadget edge” $\{v, v'\}$ simulates the effect of taking v . It is safe to remove $R := (N_2(v, w) \cap N(v)) \cup N_3(v, w)$ since, by taking v into the dominating set, all vertices in R are already dominated and since, as discussed above, it is always at least as good to take v into a minimum dominating set than to take any other of the vertices from M .

In the situation of Case (1.1), we can dominate $N_3(v, w)$ by both either v or w . Since we cannot decide at this point which of these vertices should be chosen to be in the dominating set, we use the gadget with vertices z and z' which simulates a choice between v or w , as can be seen easily. In any case, however, it is at least as good to take one of the vertices v and w (maybe both) than to take any other two vertices from M . The argument for this is similar to the one for Case (1.2). The removal of $N_3(v, w) \cup (N_2(v, w) \cap N(v) \cap N(w))$ is safe by a similar argument as the one that justified the removal of R in Case (1.2).

Finally, in Case 2, we clearly need at least two vertices to dominate $N_3(v, w)$. Since $N(v, w) \supseteq N(x, y)$ for all pairs $x, y \in M$ it is optimal to take v and w into the dominating set, simulated by the gadgets $\{v, v'\}$ and $\{w, w'\}$. As in the previous cases the removal of $N_3(v, w) \cup N_2(v, w)$ is safe since these vertices are already dominated and since these vertices need not be used for an optimal dominating set. \square

It is easy to see that applying the reduction rules to planar graphs always results in a planar graph again. This is due to the fact that the removal of vertices and edges does not affect planarity and the gadget vertices (and edges) that are introduced by Rules 1 and 2 clearly can be drawn without causing edge crossings. Here, only Case (1.1) of Rule 2 needs a little care: Since $N_3(v, w) \subseteq N(v)$ as well as $N_3(v, w) \subseteq N(w)$, the removal of $N_3(v, w)$ provides “space” for the (clearly planar) gadget drawn between v and w without any edge crossings.

Lemma 4. *Rule 2 can be carried out in time $O(n^2)$ for planar graphs and in time $O(n^4)$ for general graphs.*

Proof. To prove the time bounds for Rule 2, basically the same ideas as for Rule 1 apply (cf. proof of Lemma 2). Instead of a depth two search tree, one now has to argue on a search tree where the levels indicate the minimum of the distances to vertex v or w . Hence, we associate the vertices v and w

to the root of this search tree. The first level consists of all vertices that lie in $N(v, w)$ (i.e., at distance one from either of the vertices v or w). Determining the subset $N_1(v, w)$ means to check whether some vertex on the first level has a neighbor on the second level. We do the same kind of construction as in Lemma 2. The running time again is determined by the size of the subgraph induced by the vertices that correspond to the root and the first level of this search tree, i.e., by $G[N[v, w]]$ in this case. For planar graphs, we have $|G[N[v, w]]| = O(\deg(v) + \deg(w))$. Hence, we get $\sum_{v, w \in V} O(\deg(v) + \deg(w))$ as an upper bound on the overall running time in the case of planar graphs. Making use of the fact that $\sum_{v \in V} \deg(v) = O(n)$ for planar graphs, this is upperbounded by

$$O\left(\sum_{v, w \in V} \deg(v) + \sum_{v, w \in V} \deg(w)\right) = O(n^2).$$

In case of general graphs, we have $|G[N[v, w]]| = O((\deg(v) + \deg(w))^2)$, which trivially yields the upper bound

$$\sum_{v, w \in V} O((\deg(v) + \deg(w))^2) = O(n^4)$$

for the overall running time. □

We remark that the running times given in Lemmas 2 and 4 are pure worst-case estimates and turn out to be much lower in our experimental studies [2]. In particular, for practical purposes it is important to see that Rule 2 can only be applied for vertex pairs that are at distance at most three. The algorithms implementing these rules appear to be much faster (see the Section 4).

2.3 Reduced Graphs

We say that an application of a reduction rule leaves the graph *unchanged* if the “new” graph after applying the rule is isomorphic to the old one. Clearly, we are only interested in applications of the reduction rules that *change* the graph:

Definition 1. *Let $G = (V, E)$ be a graph such that both the application of Rule 1 and the application of Rule 2 leave the graph unchanged. Then we say that G is reduced with respect to these rules.*

Observing that the (successful) application of any reduction rule always “shrinks” the given graph implies that there can be only $O(|E|)$ successful applications of reduction rules. This leads to the following.

Theorem 1. *A graph G can be transformed into a reduced graph G' with $\gamma(G) = \gamma(G')$ in $O(n^3)$ time in the planar case and in $O(n^6)$ time in the general case.*

Proof. We prove the general statement that for a graph with m edges there can be at most $O(m)$ successful applications of reduction rules. The decisive claim we show is that after one application of Rule 1 or Rule 2 which changes the graph the resulting graph has at most the same number of vertices but at least one edge less than before the application of the rule.

Note that it is easy to verify that the total number of vertices never increases by applying the reduction rules. Now we go through Rule 1 and the various subcases of Rule 2, checking the validity of our claim. As to Rule 1, a change only occurs if there is more than one vertex affected by the rule—this means that more than one vertex and at least two edges are removed, whereas one vertex and one edge are newly introduced by the gadget.

Cases (1.2) and (1.3) of Rule 2 trivially fulfill the claim since only one gadget vertex and one gadget edge are introduced but at least two $N_3(v, w)$ vertices together with at least two incident edges are deleted. The validity of Case 2 of Rule 2 also follows easily because clearly the rule never adds more than it deletes—at least two vertices together with their edges are removed. If a change takes place, however, more edges will be removed.

Finally, concerning Case (1.1) of Rule 2 we can observe that, although the gadget introduces two more vertices and four more edges, at least the same number of vertices and more than four edges are deleted. This is true because if this case applies then at least two $N_3(v, w)$ vertices with edges to v as well as w each must exist. These and at least one additional edge will be deleted if a change takes place (otherwise, there were no change).

This concludes the proof of the claim and the theorem follows by Lemmas 2 and 4 noting that $m = O(n)$ for planar graphs and $m = O(n^2)$ for general graphs. \square

In the next section, we will make use of the following observations.

Remark 1. A graph $G = (V, E)$ which is reduced with respect to reduction Rules 1 and 2 has the following properties:

1. For all $v \in V$, the set $N_3(v)$ is always empty (these vertices are removed by Rule 1) except for it may contain a single gadget vertex of degree one.

2. For all $v, w \in V$, there exists a single vertex in $N_2(v, w) \cup N_3(v, w)$ which dominates all vertices $N_3(v, w)$ (in all other cases Rule 2 is applied).

3 A Linear Problem Kernel for Planar Graphs

Here, we show that the reduction rules given in Section 2 yield a linear size problem kernel for DOMINATING SET on planar graphs. Such a result is very unlikely to hold for general graphs, since DOMINATING SET is W[2]-complete and the existence of a (linear) problem kernel implies fixed-parameter tractability.

Theorem 2. *For a planar graph $G = (V, E)$ which is reduced with respect to Rules 1 and 2, we get $|V| \leq 335 \cdot \gamma(G)$, i.e., the DOMINATING SET problem on planar graphs admits a linear problem kernel.*

The rest of this section is devoted to the proof of Theorem 2. The proof can be split into two parts. In a first step, we try to find a so-called “maximal region decomposition” of the vertices V of a reduced graph G . In a second step, we show, on the one hand, that such a maximal region decomposition must contain all but $O(\gamma(G))$ many vertices from V . On the other hand, we prove that such a region decomposition uses at most $O(\gamma(G))$ regions, each of which containing at most $O(1)$ vertices. Combining the results then yields $|V| = O(\gamma(G))$.

The notion of “region decompositions” heavily relies on the planarity of our input graph and cannot be carried over to general graphs.

3.1 Finding a Maximal Region Decomposition

Suppose that we have a reduced planar graph G with a minimum dominating set D . We know that, in particular, neither Rule 1 applies to a vertex $v \in D$ nor Rule 2 applies to a pair of vertices $v, w \in D$. We want to get our hands on the number of vertices which lie in neighborhoods $N(v)$ for $v \in D$, or neighborhoods $N(v, w)$ for $v, w \in D$. A first idea to prove that $|V| = O(|D|)$ would be to find ($\ell = O(|D|)$ many) neighborhoods $N(v_1, w_1), \dots, N(v_\ell, w_\ell)$ with $v_i, w_i \in D$ such that all vertices in V lie in at least one such neighborhood; and then use the fact that G is reduced in order to prove that each $N(v_i, w_i)$ has size $O(1)$. Even if the graph G is reduced, however, the neighborhoods $N(v, w)$ of two vertices $v, w \in D$ may contain many vertices: the size of $N(v, w)$ in a reduced graph basically depends on how big $N_1(v, w)$ is.

In order to circumvent these difficulties, we define the concept of a region $R(v, w)$ for which we can guarantee that in a reduced graph it consists of only a constant number of vertices.

Definition 2. Let $G = (V, E)$ be a plane⁶ graph. A region $R(v, w)$ between two vertices v, w is a closed subset of the plane with the following properties:

1. the boundary of $R(v, w)$ is formed by two simple paths P_1 and P_2 in V which connect v and w , and the length of each path is at most three⁷, and
2. all vertices which are strictly inside⁸ the region $R(v, w)$ are from $N(v, w)$.

For a region $R = R(v, w)$, let $V(R)$ denote the vertices belonging to R , i.e.,

$$V(R) := \{u \in V \mid u \text{ sits inside or on the boundary of } R\}.$$

In the following, the boundary of a region R will be denoted by ∂R .

Definition 3. Let $G = (V, E)$ be a plane graph and $D \subseteq V$. A D -region decomposition of G is a set \mathcal{R} of regions between pairs of vertices in D such that

1. for $R(v, w) \in \mathcal{R}$ no vertex from D (except for v, w) lies in $V(R(v, w))$ and
2. for two regions $R_1, R_2 \in \mathcal{R}$, it holds $(R_1 \cap R_2) \subseteq (\partial R_1 \cup \partial R_2)$.

For a D -region decomposition \mathcal{R} , we define $V(\mathcal{R}) := \bigcup_{R \in \mathcal{R}} V(R)$. A D -region decomposition \mathcal{R} is called maximal if there is no region $R \notin \mathcal{R}$ such that $\mathcal{R}' := \mathcal{R} \cup \{R\}$ is a D -region decomposition where $V(\mathcal{R})$ is a strict subset of $V(\mathcal{R}')$.

For an example of a (maximal) D -region decomposition we refer to the left-hand side diagram of Fig. 3.

We will show that, for a given graph G with dominating set D , we can always find a maximal D -region decomposition with at most $O(\gamma(G))$ many regions. For that purpose, we observe that a D -region decomposition induces a graph in a very natural way.

⁶A plane graph is a particular planar embedding of a planar graph.

⁷The length of a path is the number of edges on it.

⁸By “strictly inside the region $R(v, w)$ ” we mean lying in the region, but not sitting on the boundary of $R(v, w)$.

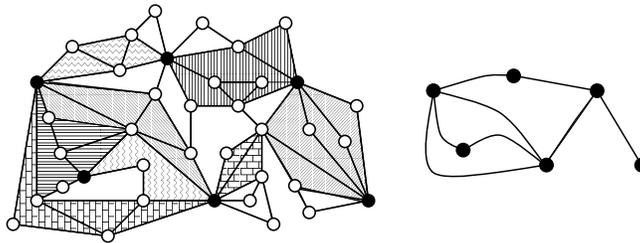


Figure 3: The left-hand side diagram shows an example of a possible D -region decomposition \mathcal{R} of some graph G , where D is the subset of vertices in G that are drawn in black. The various regions are highlighted by different patterns. The remaining white areas are not considered as regions. The given D -region decomposition is maximal. The right-hand side shows the induced graph $G_{\mathcal{R}}$ (Definition 4).

Definition 4. *The induced graph $G_{\mathcal{R}} = (V_{\mathcal{R}}, E_{\mathcal{R}})$ of a D -region decomposition \mathcal{R} of G is the graph with possible multiple edges which is defined by $V_{\mathcal{R}} := D$ and*

$$E_{\mathcal{R}} := \{\{v, w\} \mid \text{there is a region } R(v, w) \in \mathcal{R} \text{ between } v, w \in D\}.$$

Note that, by Definition 3, the induced graph $G_{\mathcal{R}}$ of a D -region decomposition is planar. For an example of an induced graph $G_{\mathcal{R}}$ see Fig. 3.

Definition 5. *A planar graph $G = (V, E)$ with multiple edges is thin if there exists a planar embedding such that no two multiedges are homotopic: This means that if there are two edges e_1, e_2 between a pair of distinct vertices $v, w \in V$, then there must be two further vertices $u_1, u_2 \in V$ which sit inside the two disjoint areas of the plane that are enclosed by e_1, e_2 .*

The induced graph $G_{\mathcal{R}}$ in Fig. 3 is thin.

Lemma 5. *For a thin planar graph $G = (V, E)$ we have $|E| \leq 3|V| - 6$.*

Proof. The claim is true for planar graphs without multiple edges. We prove the claim by an induction on the number ℓ_G of multiple edges in G . More precisely, for a graph $G = (V, E)$ with multiple edges (i.e., E is a multiset), we let

$$\ell_G := \frac{1}{2} \left(\sum_{v, w \in V} \left(\sum_{\{v, w\} \in E} 1 \right) - 1 \right)$$

For $\ell_G = 0$, the claim is true, since a planar graph (without multiple edges) has at most $3|V| - 6$ edges. Now, suppose the claim is true for all graphs which have at most ℓ_G multiple edges. Consider a planar graph $G = (V, E)$ with $\ell_G + 1$ multiple edges. Choose a pair of vertices $v, w \in V$ which is connected by at least two edges $e_1, e_2 \in E$. Since G is thin, we may consider a planar embedding, in which e_1 and e_2 are not homotopic. Let $G_1 = (V_1, E_1)$ be the subgraph of G which consists of the vertices v, w , the edge e_1 and all vertices and edges that sit strictly inside the area A of the plane that is enclosed by e_1 and e_2 . Similarly, let $G_2 = (V_2, E_2)$ be the subgraph of G which consists of the vertices v, w , the edge e_2 and all vertices and edges that sit strictly outside the area A . Hence, we have $|E| = |E_1| + |E_2|$ and $|V| = |V_1| + |V_2| - 2$. Since, by construction, $\ell_{G_1}, \ell_{G_2} < \ell_G$, the induction hypothesis yields

$$\begin{aligned} |E| &= |E_1| + |E_2| \\ &\leq (3|V_1| - 6) + (3|V_2| - 6) \\ &= 3|V| - 6. \end{aligned}$$

□

Using the notion of thin graphs, we can formulate the main result of this subsection.

Proposition 1. *For a reduced plane graph G with dominating set D , there exists a maximal D -region decomposition \mathcal{R} such that $G_{\mathcal{R}}$ is thin.*

Proof. We give a constructive proof on how to find a maximal D -region decomposition \mathcal{R} of a plane graph G such that the induced graph $G_{\mathcal{R}}$ is thin. Consider the algorithm presented in Fig. 4. It is obvious that the algorithm returns a D -region decomposition, since—by construction—we made sure that regions are between vertices in D , that regions do not contain vertices from D , and that regions do not intersect. Moreover, the D -region decomposition obtained by the algorithm is maximal: If a vertex u does not belong to a region, i.e., if $u \notin V_{\text{used}}$, then the algorithm eventually checks, whether there is a region S_u such that $\mathcal{R} \cup \{S_u\}$ is a D -region decomposition. It remains to show that the induced graph $G_{\mathcal{R}}$ of the D -region decomposition \mathcal{R} found by the algorithm is thin. We embed $G_{\mathcal{R}}$ in the plane in such a way that an edge belonging to a region $R \in \mathcal{R}$ is drawn inside the area covered by R . To see that the graph is thin, we have to show that, for every multiple edge e_1, e_2 (belonging to two regions $R_1, R_2 \in \mathcal{R}$ that were chosen

```

region_decomp(plane graph  $G = (V, E)$ , vertex subset  $D \subseteq V$ )
// Returns a  $D$ -region decomposition  $\mathcal{R}$  for  $G$  such that
// the induced graph  $G_{\mathcal{R}}$  is thin.

```

- Let $V_{\text{used}} \leftarrow \emptyset$; $\mathcal{R} \leftarrow \emptyset$.
- For all $u \in V$ do
 - If ($(u \notin V_{\text{used}})$ and ($u \in V(R)$ for some region $R = R(v, w)$ between two vertices $v, w \in D$ such that $\mathcal{R} \cup \{R\}$ is a D -region decomposition)) then
 - * Consider the set \mathcal{R}_u of all regions S with the following properties:^a
 1. S is a region between v and w .
 2. S contains u .
 3. no vertex from $D \setminus \{v, w\}$ is in $V(S)$.
 4. S does not cross any region from \mathcal{R} , i.e., $(S \cap R) \subseteq (\partial S \cup \partial R)$ for all $R \in \mathcal{R}$.
 - * Choose a region $S_u \in \mathcal{R}_u$ which is maximal in space.^b
 - * $\mathcal{R} \leftarrow \mathcal{R} \cup \{S_u\}$.
 - * $V_{\text{used}} \leftarrow V_{\text{used}} \cup V(S_u)$.
- Return \mathcal{R} .

^aThese four properties ensure that $\mathcal{R} \cup \{S\}$ is a D -region decomposition for every $S \in \mathcal{R}_u$.

^bA region S_u is maximal in space if $S' \supseteq S_u$ for any $S' \in \mathcal{R}_u$ implies $S' = S_u$.

Figure 4: Greedy-like construction of a maximal D -region decomposition.

at some point of the algorithm) between two vertices $v, w \in D$, there exist two vertices $u_1, u_2 \in D$ which lie inside the areas enclosed by e_1, e_2 . Let A be such an area. Suppose that there is no vertex $u \in D$ in A . We distinguish two cases. Either there is also no vertex from $V \setminus D$ in A or there are other vertices V' from $V \setminus D$ inside A . In the first case, by joining the regions R_1 and R_2 we obtain a bigger region which fulfills all the four conditions checked by the algorithm in Fig. 4, a contradiction to the maximality of R_1 and R_2 . In the second case, since D is assumed to be a dominating set, the vertices in V' need to be dominated by D . Since v, w are the only vertices from D which are part of A , R_1 or R_2 , the vertices in V' need to be dominated by v, w , hence they belong to $N(v, w)$. But then again by joining the regions R_1 and R_2 we obtain a bigger region which again fulfills all the four conditions of the algorithm in Fig. 4, a contradiction to the maximality of R_1 and R_2 . \square

3.2 Region Decompositions and the Size of Reduced Planar Graphs

Suppose that we are given a reduced plane graph $G = (V, E)$ with a minimum dominating set D . Then, by Proposition 1 and Lemma 5, we can find a maximal D -region decomposition \mathcal{R} of G with at most $O(\gamma(G))$ regions. In order to see that $|V| = O(\gamma(G))$, it remains to show that

1. there are at most $O(\gamma(G))$ vertices which do not belong to any of the regions in \mathcal{R} , and that
2. every region of \mathcal{R} contains at most $O(1)$ vertices.

These issues are treated by the following two propositions.

We first of all state two technical lemmas, one which characterizes an important property of a maximal region decomposition and another one which gives an upper bound on the size of a special type of a region.

Lemma 6. *Let G be a reduced plane graph with a dominating set D and let \mathcal{R} be a maximal D -region decomposition. If $u \in N_1(v)$ for some vertex $v \in D$ then $u \in V(\mathcal{R})$.*

Proof. In the following, we say that an edge *crosses* a region R , if the edge lies (possibly except for its endpoints) strictly inside R . Similarly we say that a path *crosses* a region R if at least one edge of the path crosses R .

Let $u \in N_1(v)$ for some $v \in D$ and assume that $u \notin V(\mathcal{R})$. By definition of $N_1(v)$, there exists a vertex $u' \in N(u)$ with $u' \notin N[v]$. We distinguish two cases. Either $u' \in D$ or u' needs to be dominated by a vertex $w \in D$ with $w \neq v$. If $u' \in D$, we consider the (degenerated) region consisting of the path $\{v, u, u'\}$. Since \mathcal{R} is assumed to be maximal, this path must cross a region $R \in \mathcal{R}$. But this implies that $u \in V(R)$, a contradiction.

In the second case, we consider the (degenerated) region consisting of the path $\langle v, u, u', w \rangle$. Again, by maximality of \mathcal{R} , this path must cross a region $R = R(x, y) \in \mathcal{R}$ between two vertices $x, y \in D$. Since, by assumption, $u \notin V(R)$, neither the edge $\{v, u\}$, nor the edge $\{u, u'\}$ can cross R . This implies that the edge $\{u', w\}$ crosses R . From this we know that w lies on the boundary of or inside R and, hence, $w \in V(R)$. However, according to the definition of a D -region decomposition, the only vertices from D that are in $V(R)$ are x, y . Hence, w.l.o.g., $x = w$. At the same time u' must lie on the boundary of R , otherwise $u \in V(R)$. By definition of a region, there exists path P of length at most three between w and y that goes through u' and that is part of the boundary of R . Observe that $u' \neq y$, since $y \in D$ and

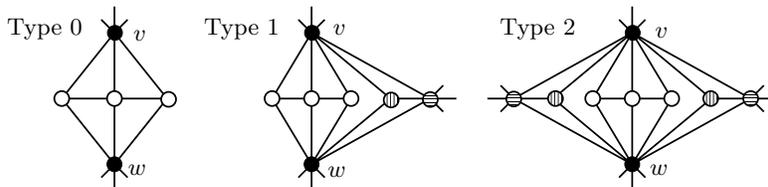


Figure 5: Simple regions of Type 0, Type 1, Type 2. This figure illustrates the largest possible simple regions in a reduced graph. Vertices marked with horizontal lines are in $N_1(v, w)$, vertices marked with vertical lines belong to $N_2(v, w)$, and white vertices are in $N_3(v, w)$.

we assume that $u' \notin D$. We claim, however, that u' is a neighbor of y : To see this, observe that, the edge $\{w, u'\}$ cannot be part of P , since we already know that this edge crosses R . As a consequence, the path P uses more than one edge in order to reach u' from w . On the other hand, since $u' \neq y$, and P has length at most three, we know that the path P (between w and y) uses exactly two edges to reach u' from w . This, however, implies that u' is a neighbor of y as claimed. But then, the (degenerated) region R' consisting of the path $\{v, u, u', y\}$ is a region between two vertices v and y in D , which does not cross (it only touches R) any region in \mathcal{R} . For the D -region decomposition $\mathcal{R}' := \mathcal{R} \cup \{R'\}$, we have $u \in V(\mathcal{R}') \setminus V(\mathcal{R})$, contradicting the maximality of \mathcal{R} . \square

We now investigate a special type of a region specified by the following definition.

Definition 6. A region $R(v, w)$ between two vertices $v, w \in D$ is called simple if all vertices contained in $R(v, w)$ except for v, w are common neighbors of both v and w , i.e., if $(V(R(v, w)) \setminus \{v, w\}) \subseteq N(v) \cap N(w)$.

Let v, u_1, w, u_2 be the vertices that sit on the boundary of the simple region $R(v, w)$, when walking along the boundary in clockwise order. We say that $R(v, w)$ is a simple region of Type i ($0 \leq i \leq 2$) if i vertices from $\{u_1, u_2\}$ have a neighbor outside $R(v, w)$.

Lemma 7. Every simple region R of Type i of a plane reduced graph contains at most $5 + 2i$ vertices.

Proof. Let $R = R(v, w)$ be a simple region of Type i between vertices v and w . We will show that $|V(R)| \leq 5 + 2i$. The worst-case simple regions

are depicted in Fig. 5. Firstly, let us count the number of vertices in $V(R)$ which belong to $N_1(v, w) \cup N_2(v, w)$. Clearly, only vertices on the boundary (except for v and w) can have a neighbor outside R . Thus, all vertices in $N_1(v, w) \cap V(R)$ lie on the boundary of R . By definition of a simple region of Type i , we have $|N_1(v, w) \cap V(R)| \leq i$. Moreover, it is easy to see that, by planarity, every vertex in $N_1(v, w) \cap V(R)$ can contribute at most one vertex to $N_2(v, w) \cap V(R)$. Hence, we get $|(N_1(v, w) \cup N_2(v, w)) \cap V(R)| \leq 2i$. Secondly, we determine the number of vertices in $N_3(v, w) \cap V(R)$. Since G is reduced, by Remark 1, we know that these vertices need to be dominated by a single vertex in $N_2(v, w) \cup N_3(v, w)$. Moreover, since the region is simple, all vertices in $N_3(v, w) \cap V(R)$ are neighbors of both v and w . By planarity, it follows that there can be at most 3 vertices in $N_3(v, w) \cap V(R)$. In summary, together with the vertices $v, w \in V(R)$, we get $|V(R)| \leq 5 + 2i$. \square

We use Lemmas 6 and 7 for the following two proofs.

Proposition 2. *Let $G = (V, E)$ be a plane reduced graph and let D be a dominating set of G . If \mathcal{R} is a maximal D -region decomposition then $|V \setminus V(\mathcal{R})| \leq 2|D| + 56|\mathcal{R}|$.*

Proof. We claim that every vertex $u \in V \setminus V(\mathcal{R})$ is either a vertex in D or belongs to a set $N_2(v) \cup N_3(v)$ for some $v \in D$. To see this, suppose that $u \notin D$. But since D is a dominating set, we know that $u \in N(v) = N_1(v) \cup N_2(v) \cup N_3(v)$ for some vertex $v \in D$. Since \mathcal{R} is assumed to be maximal, by Lemma 6, we know that $N_1(v) \subseteq V(\mathcal{R})$. Thus, $u \in N_2(v) \cup N_3(v)$.

For a vertex $v \in D$, let $N_2^*(v) = N_2(v) \setminus V(\mathcal{R})$. The above observation implies that $V \setminus V(\mathcal{R}) \subseteq D \cup (\bigcup_{v \in D} N_3(v)) \cup (\bigcup_{v \in D} N_2^*(v))$.

We, firstly, upperbound the size of $\bigcup_{v \in D} N_3(v)$. Since, by Remark 1, $|N_3(v)| \leq 1$, we get $|\bigcup_{v \in D} N_3(v)| \leq |D|$.

We now upperbound the size of $N_2^*(v)$ for a given vertex $v \in D$. To this end, for a vertex $v \in D$, let $N_1^*(v)$ be the subset of $N_1(v)$ which sit on the boundary of a region in \mathcal{R} . It is clear that $N_2^*(v) \subseteq N(v) \cap N(N_1^*(v))$. Hence, we investigate the set $N_1^*(v)$. Suppose that $R(v, w_1), \dots, R(v, w_\ell)$ are the regions between v and some other vertices $w_i \in D$, where $\ell = \deg_{G_{\mathcal{R}}}(v)$ is the degree of v in the induced region graph $G_{\mathcal{R}}$. Then, every region $R(v, w_i)$ can contribute at most two vertices u_i^1, u_i^2 to $N_1^*(v)$, i.e., in the worst-case, we have $N_1^*(v) = \bigcup_{i=1}^{\ell} \{u_i^1, u_i^2\}$ with $u_i^1, u_i^2 \in V(R(v, w_i))$, i.e., $|N_1^*(v)| \leq 2 \deg_{G_{\mathcal{R}}}(v)$. We already observed that every vertex in $N_2^*(v)$ must be a common neighbor of v and some vertex in $N_1^*(v)$. We claim that, moreover,

the vertices in $N_2^*(v)$ can be grouped into various simple regions. More precisely, we claim that there exists a set \mathcal{S}_v of simple regions such that

1. every $S \in \mathcal{S}_v$ is a simple region between v and some vertex in $N_1^*(v)$,
2. $N_2^*(v) \subseteq \bigcup_{S \in \mathcal{S}_v} V(S)$, and
3. $|\mathcal{S}_v| \leq 2 \cdot |N_1^*(v)|$.

The idea for the construction of the set \mathcal{S}_v is similar to the greedy-like construction of a maximal region decomposition (see Fig. 4). Starting with \mathcal{S}_v as empty set, one iteratively adds a *simple* region $S(v, x)$ between v and some vertex $x \in N_1^*(v)$ to the set \mathcal{S}_v in such a way that (1) $\mathcal{S}_v \cup \{S(v, x)\}$ contains more $N_2^*(v)$ -vertices than \mathcal{S}_v , (2) $S(v, x)$ does not cross any region in \mathcal{S}_v and (3) $S(v, x)$ is maximal (in space) under all simple regions S between v and x that do not cross any region in \mathcal{S}_v . The fact that we end up with at most $2 \cdot |N_1^*(v)|$ many regions can be seen as follows. Consider the induced graph $G_{\mathcal{S}_v}$, which has the set $\{v\} \cup N_1^*(v)$ as vertices and an edge between v and a vertex $u \in N_1^*(v)$ if and only if \mathcal{S}_v contains a simple region between v and u . In other words, $G_{\mathcal{S}_v}$ is a star with possible multiple edges. Since, by construction, all simple regions were chosen maximal in space, the graph $G_{\mathcal{S}_v}$ is thin. It is not hard to see that a thin star on $n + 1$ vertices can have at most $2n$ edges. In particular, this shows that $G_{\mathcal{S}_v}$ has at most $2 \cdot |N_1^*(v)|$ edges, i.e., $|\mathcal{S}_v| \leq 2 \cdot |N_1^*(v)|$.

Since, by Lemma 7, every simple region $S(v, x)$ with $x \in N_1^*(v)$ contains at most seven vertices—not counting the vertices v and x which clearly cannot be in $N_2^*(v)$ —we conclude that $|N_2^*(v)| \leq 7 \cdot |\mathcal{S}_v| \leq 14 \cdot |N_1^*(v)| \leq 28 \cdot \deg_{G_{\mathcal{R}}}(v)$. From the fact that $V \setminus V(\mathcal{R}) \subseteq D \cup (\bigcup_{v \in D} N_3(v)) \cup (\bigcup_{v \in D} N_2^*(v))$ (see above) we then get

$$|V \setminus V(\mathcal{R})| \leq |D| + |D| + \sum_{v \in D} |N_2^*(v)| \leq 2 \cdot |D| + 28 \sum_{v \in D} \deg_{G_{\mathcal{R}}}(v) \leq 2 \cdot |D| + 56 \cdot |\mathcal{R}|.$$

□

We now investigate the maximal size of a region in a reduced graph. The worst-case scenario for a region in a reduced graph is depicted in Fig. 6.

Proposition 3. *A region R of a plane reduced graph contains at most 55 vertices, i.e., $|V(R)| \leq 55$.*

Proof. Let $R = R(v, w)$ be a region between vertices $v, w \in V$. As in the proof of Lemma 7, we count the number of vertices in $V(R) \subseteq N[v, w]$ which belong to $N_1(v, w)$, $N_2(v, w)$, and $N_3(v, w)$, separately.

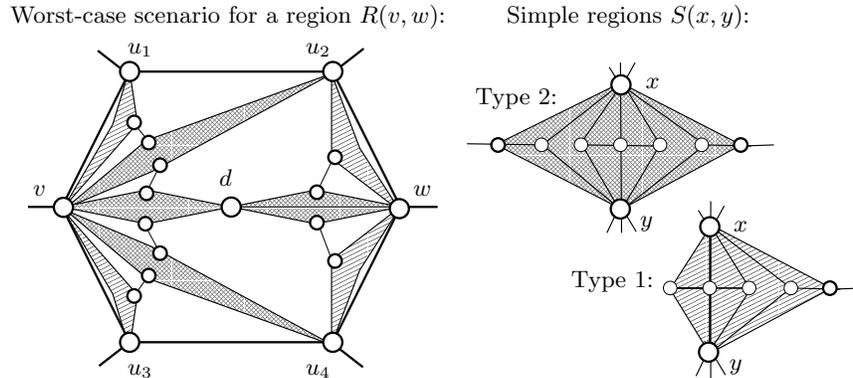


Figure 6: The left-hand diagram shows a worst-case scenario for a region $R(v, w)$ between two vertices v and w in a reduced planar graph (cf. the proof of Proposition 3). Such a region may contain up to four vertices from $N_1(v, w)$, namely u_1, u_2, u_3 , and u_4 . The vertices from $R(v, w)$ which belong to the sets $N_2(v, w)$ and $N_3(v, w)$ can be grouped into so-called simple regions of Type 1 (marked with a line-pattern) or of Type 2 (marked with a crossing-pattern); the structure of such simple regions $S(x, y)$ is given in the right-hand part of the diagram. In $R(v, w)$ there might be two simple regions $S(d, v)$ and $S(d, w)$ (of Type 2), containing vertices from $N_3(v, w)$. And, we can have up to six simple regions of vertices from $N_2(v, w)$: $S(u_1, v), S(v, u_3), S(u_4, w), S(w, u_2), S(u_2, v)$, and $S(u_4, v)$ (among these, the latter two can be of Type 2 and the others are of Type 1). See the proof of Proposition 3 for details.

We start with the number of vertices in $N_3(v, w) \cap V(R)$. Since the graph is assumed to be reduced, by Remark 1, we know that all vertices in $N_3(v, w)$ need to be dominated by a single vertex from $N_2(v, w) \cup N_3(v, w)$. Denote by d the vertex which dominates all vertices in $N_3(v, w)$. Since all vertices in $N_3(v, w)$ are also dominated by v or w , we may write $N_3(v, w) = S(d, v) \cup S(d, w)$ where $S(d, v) \subseteq N(d) \cap N(v)$ and $S(d, w) \subseteq N(d) \cap N(w)$. In this way, $S(d, v)$ and $S(d, w)$ form simple regions between d and v , and d and w , respectively. In Fig. 6 these simple regions $S(d, v)$ and $S(d, w)$ (of Type 2) are drawn with a crossing pattern. By Lemma 7 we know that $S(d, v)$ and $S(d, w)$ both contain at most seven vertices each, not counting the vertices d, v and d, w , respectively. Since d maybe from $N_3(v, w)$, we obtain $|N_3(v, w) \cap V(R)| \leq 2 \cdot 7 + 1 = 15$.

It is clear that vertices in $N_1(v, w) \cap V(R)$ need to be on the boundary of R ,

since, by definition of $N_1(v, w)$, they have a neighbor outside $N(v, w)$. The region R is enclosed by two paths P_1 and P_2 between v and w of length at most three each. Hence, there can be at most four vertices in $N_1(v, w) \cap V(R)$, where this worst-case holds if P_1 and P_2 are disjoint and have length exactly three each. Consider Fig. 6, which shows a region enclosed by two such paths. Suppose that the four vertices on the boundary besides v and w are u_1, u_2, u_3 , and u_4 .

Finally, we count the number of vertices in $N_2(v, w) \cap V(R)$. It is important to note that, by definition of $N_2(v, w)$, every such vertex needs to have a neighbor in $N_1(v, w)$ and at the same time needs to be a neighbor of either v or w (or both). Hence, $N_2(v, w) = \bigcup_{i=1}^4 (S(u_i, v) \cup S(u_i, w))$, where $S(u_i, v) \subseteq N(u_i) \cap N(v)$ and $S(u_i, w) \subseteq N(u_i) \cap N(w)$. All the sets $S(u_i, v)$ and $S(u_i, w)$, where $1 \leq i \leq 4$, form simple regions inside R . Due to planarity, however, there cannot exist all eight of these regions. In fact, in order to avoid crossings, the worst-case scenario is depicted in Fig. 6 where six of these simple regions exist (they are drawn with a line-pattern in the figure).⁹ Concerning the type of these simple regions, it is not hard to verify, that in the worst-case there can be two among these six regions of Type 2, the other four of them being of Type 1. In Fig. 6, the simple regions $S(u_2, v)$ and $S(u_4, v)$ are of Type 2 (having two connections to vertices outside the simple region), and the simple regions $S(u_1, v), S(u_2, w), S(u_3, v)$, and $S(u_4, w)$ are of Type 1 (having only one connection to vertices outside the region; a second connection to vertices outside the region is not possible because of the edges $\{u_1, v\}, \{u_2, w\}, \{u_3, v\}$, and $\{u_4, w\}$). In summary, the worst-case number of vertices in $N_2(v, w) \cap V(R)$ is given by four times the number of vertices of a simple region of Type 1 and two times the number of vertices of a simple region of Type 2; each time, of course, excluding vertices from $\{u_1, u_2, u_3, u_4, v, w\}$. By Lemma 7 this amounts to $|N_2(v, w) \cap V(R)| \leq 4 \cdot (3 + 2 \cdot 1) + 2 \cdot (3 + 2 \cdot 2) = 34$.¹⁰

The claim now follows from the fact that $V(R) = \{v, w\} \cup (V(R) \cap N_3(v, w)) \cup (V(R) \cap N_1(v, w)) \cup (V(R) \cap N_2(v, w))$, which yields $|V(R)| = 2 + 15 + 4 + 34 = 55$. \square

In summary, in order to prove Theorem 2 we first of all observe that, for a graph G with minimum dominating set D , by Proposition 1 and Lemma 5, we can find a D -region decomposition \mathcal{R} of G with at most $3\gamma(G)$ regions,

⁹Observe that regions $S(u_1, w)$ and $S(u_3, w)$ would cross the regions $S(u_2, v)$ and $S(u_4, v)$, respectively.

¹⁰Note that for the size of, e.g., a region $S(u_i, v)$ we do not have to count u_i and v , since they are not vertices in $N_2(v, w)$.

i.e., $|\mathcal{R}| \leq 3\gamma(G)$. By Proposition 3, we know that $|V(\mathcal{R})| \leq \sum_{R \in \mathcal{R}} |V(R)| \leq 55|\mathcal{R}|$. By Proposition 2, we have $|V \setminus V(\mathcal{R})| \leq 2|D| + 56|\mathcal{R}|$. Hence, we get $|V| \leq 2|D| + 111|\mathcal{R}| \leq 335\gamma(G)$.

4 Concluding Remarks

In this work, two lines of research meet. On the one hand, there is **DOMINATING SET**, one of the NP-complete core problems of combinatorial optimization and graph theory. On the other hand, the second line of research is that of algorithm engineering and, in particular, the power of data reduction by efficient preprocessing. Presenting two simple and easy to implement reduction rules for **DOMINATING SET**, we proved that for planar graphs a linear size problem kernel can be efficiently constructed. Our result complements and partially improves previous results [3, 4, 5, 6, 25, 32] on the parameterized complexity of **DOMINATING SET** on planar graphs. We emphasize that the proven bound on the problem kernel size is a pure worst-case upper bound. In practice, we obtained much smaller problem kernels (see below). An immediate open question is to further lower the worst-case upper bound on the size of the problem kernel, improving the constant factor to values say around 10. This would bring the problem kernel for **DOMINATING SET** on planar graphs into “dimensions” as known for **VERTEX COVER**, where it is of “optimal” size $2k$ [13]. This could be done by either improving the analysis given or (more importantly) further improving the given reduction rules or both. Improving the rules might be done by further extending the concept of neighborhood to more than two vertices. From a practical point of view, however, one also has to take into account to keep the reduction rules as simple as possible in order to avoid inefficiency due to increased overhead. It might well be the case that additional, more complicated reduction rules only improve the worst-case bounds, but are of little or no practical use due to their computational overhead. A question that deserves further attention, however, is to find out whether by the use of dynamic graph data structures or other implementation tricks the worst-case time complexity of our rules can be significantly improved.

It might be interesting to see whether similar reduction rules with a provable guarantee on the size of the reduced instances can also be found for variations of **DOMINATING SET** problem, such as **TOTAL DOMINATING SET**, or **PERFECT DOMINATING SET** (see [39] for a description of such variants). The study of preprocessing by reduction rules is valuable for various other problems (see [24] for a recent survey).

Finally, we mention that the techniques in this paper are of a topological nature and might carry over to prove a similar result (including, however, the genus into the linear size factor for the problem kernel) for DOMINATING SET on graphs of bounded genus. Recently, there has been increased interest in solving domination-like problems on somewhat more general graph classes than planar ones—cf., e.g., [14, 15, 16, 20, 25, 26]. In particular, an open question is whether a linear problem kernel can also be proven for other graph classes such as, e.g., disk intersection graphs, for which the parameterized complexity of DOMINATING SET is not known (see [7]). Altogether, we would like to emphasize that basically all the cited work on domination-like problems on planar and related graphs seems to be of purely theoretical nature with so far no impact in practical computing. By way of contrast, our work delivers easy to implement reduction rules whose value has been proven in experimental work [2].

Experimental studies. We briefly report on the efficiency of the given reduction rules in some experiments with random planar graphs. More experimental results in particular with respect to “Internet graphs” can be found in [2]. The performance of the preprocessing was measured on a set of combinatorial random planar graphs of various sizes. These graphs have been generated with the standard function provided by the algorithm library LEDA [34].¹¹ More precisely, we created eight sample sets of 100 random planar graphs each, containing instances with 100, 500, 750, 1000, 1500, 2000, 3000, and 4000 vertices. The preprocessing seems, at least on the given random sample sets, to be very effective. As a general rule of thumb, we may say that, in all of the cases,

- more than 79% of the vertices and
- more than 88% of the edges

were removed from the graph. Moreover, the reduction rules determined a very high percentage (for all cases approximately 89%) of the vertices of an optimal dominating set. The overall running time for the reduction ranged from less than one second (for small graph instances with 100 vertices) to around 30 seconds (for larger graph instances with 4000 vertices).

We remark that, in our experiments, we used a slight modification of the reduction rules: Formally, when Rule 1 or Rule 2 is applied and some vertex v

¹¹For each instance with n vertices, first a “maximal planar graph” with $3n - 6$ edges is randomly generated, then a number m between $n - 1$ and $3n - 6$ is randomly chosen and all but m edges are removed from the graph. We remark that this method does not generate graphs according to the uniform distribution (see [34] for details).

is determined to belong to an optimal dominating, the reduction rules attach a gadget vertex v' of degree one to v . In our setting, we simply removed the vertex v from the graph and “marked” its neighbors as being already dominated. In this sense, we dealt with an annotated version of DOMINATING SET, where the input instances are black-and-white graphs consisting of two types of vertices: black vertices which still need to be dominated; and white vertices which are assumed to be already dominated. A slight modification makes Rule 1 and Rule 2 applicable to such instances as well.

Finally, we enriched our reduction rules by further heuristics. We additionally used three (very simple) extra rules that were presented in the search tree algorithm in [4]. These extra rules are concerned with the removal of white vertices in such black-and-white graphs for the annotated version of DOMINATING SET (for the details and their correctness see [4]): (1) delete a white vertex of degree zero or one; (2) delete a white vertex of degree two if its neighbors are at distance at most two from each other; (3) delete a white vertex of degree three if the subgraph induced by its neighbors is connected. Enriching our reduction rules with these extra rules led to a very powerful data reduction on our set of random instances described above. We observed that in this extended setting, the running times for the data reduction went down to less than half a second (for graphs of 100 vertices) and less than eight seconds (for graphs of 4000 vertices) in average. Most interestingly, the combination of these rules removed, in average,

- more than 99.7% of the vertices and
- more than 99.8% of the edges

of the original graph. A similarly high percentage of the vertices that belong to an optimal dominating set could be detected. A more thorough discussion of the experiments with random planar graphs can be found in [1] and experiments with “Internet graphs” (which are sparse but not planar) can be found in [2].

Acknowledgements. For two years, besides ourselves the linear size problem kernel question for DOMINATING SET on planar graphs has taken the attention of numerous people, all of whom we owe sincere thanks for their insightful and inspiring remarks and ideas. Among these people we particularly would like to mention Nadja Betzler, Britta Dorn, Frederic Dorn, Henning Fernau, Jens Gramm, Michael Kaufmann, Ton Kloks, Klaus Reinhardt, Fran Rosamond, Peter Rossmanith, Ulrike Stege, and Pascal Tesson. Special thanks go to Henning for the many hours he spent with us on “diamond discussions” and for pointing us to a small error concerning the application of the linear problem kernel and to Frederic for implementing the rules, which also uncovered a small error in a previous version of reduction Rule 2.

We are grateful to an anonymous referee for comments that helped improve the presentation of the paper.

References

- [1] J. Alber. *Exact Algorithms for NP-hard Problems on Networks: Design, Analysis, and Implementation*. PhD thesis, Universität Tübingen, Germany. January 2003.
- [2] J. Alber, N. Betzler, and R. Niedermeier. Experiments on data reduction for optimal domination in networks. To appear in *Proceedings International Network Optimization Conference (INOC 2003)*, Evry/Paris, France, October 2003.
- [3] J. Alber, H. L. Bodlaender, H. Fernau, T. Kloks, and R. Niedermeier. Fixed parameter algorithms for dominating set and related problems on planar graphs. *Algorithmica*, **33**(4): 461–493, 2002.
- [4] J. Alber, H. Fan, M. R. Fellows, H. Fernau, R. Niedermeier, F. Rosamond, and U. Stege. Refined search tree technique for DOMINATING SET on planar graphs. In *Proceedings 26th MFCS 2001*, Springer-Verlag LNCS 2136, pp. 111–122, 2001.
- [5] J. Alber, H. Fernau, and R. Niedermeier. Parameterized complexity: exponential speed-up for planar graph problems. In *Proceedings 28th ICALP 2001*, Springer-Verlag LNCS 2076, pp. 261–272, 2001.
- [6] J. Alber, H. Fernau, and R. Niedermeier. Graph separators: a parameterized view. In *Proceedings 7th COCOON 2001*, Springer-Verlag LNCS 2108, pp. 318–327, 2001. Long version to appear in *Journal of Computer and System Sciences*.
- [7] J. Alber and J. Fiala. Geometric separation and exact solutions for the parameterized independent set problem on disk graphs. In *Proceedings 2nd IFIP TCS 2002*, Kluwer Academic Press, pp. 26–37, 2002.
- [8] J. Alber, J. Gramm, and R. Niedermeier. Faster exact solutions for hard problems: a parameterized point of view. *Discrete Mathematics*, **229**: 3–27, 2001.
- [9] J. Alber and R. Niedermeier. Improved tree decomposition based algorithms for domination-like problems. In *Proceedings 5th LATIN 2002*, Springer-Verlag LNCS 2286, pp. 613–627, 2002.
- [10] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi. *Complexity and Approximation*. Springer-Verlag, 1999.

- [11] B. S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *Journal of the ACM*, **41**:153–180, 1994.
- [12] R. Bar-Yehuda and S. Even. A local-ratio theorem for approximating the weighted vertex cover problem. *Annals of Discrete Mathematics*, **25**: 27–46, 1985.
- [13] J. Chen, I. A. Kanj, and W. Jia. Vertex cover: further observations and further improvements. *Journal of Algorithms*, **41**:280–301, 2001.
- [14] J. Chen, I. A. Kanj, L. Perkovic, E. Sedgwick, and G. Xia. Genus characterizes the complexity of graph problems: some tight results. In *Proceedings 30th ICALP 2003*, Springer-Verlag LNCS 2719, pp. 845–856, 2003.
- [15] E. D. Demaine, Fedor V. Fomin, M. Taghi Hajiaghayi, and D. M. Thilikos. Fixed-parameter algorithms for the (k, r) -center in planar graphs and map graphs. In *Proceedings 30th ICALP 2003*, Springer-Verlag LNCS 2719, pp. 829–844, 2003.
- [16] E. D. Demaine, M. Taghi Hajiaghayi, and D. M. Thilikos. Exponential speedup of fixed-parameter algorithms on $K_{3,3}$ -minor-free or K_5 -minor-free graphs. In *Proceedings 13th ISAAC 2002*, Springer-Verlag LNCS 2518, pp. 262–273, 2002.
- [17] R. G. Downey. Parameterized complexity for the skeptic (invited paper). In *Proc. of 18th IEEE Conference on Computational Complexity*, pp. 147–168, 2003.
- [18] R. G. Downey and M. R. Fellows. Fixed-parameter tractability and completeness. *Congressus Numerantium*, **87**:161–187, 1992.
- [19] R. G. Downey and M. R. Fellows. *Parameterized Complexity*. Monographs in Computer Science. Springer-Verlag, 1999.
- [20] J. Ellis, H. Fan, and M. R. Fellows. The dominating set problem is fixed parameter tractable for graphs of bounded genus. In *Proceedings 8th SWAT 2002*, Springer-Verlag LNCS 2368, pp. 180–189, 2002.
- [21] U. Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, **45**:634–652, 1998.
- [22] M. R. Fellows. Parameterized complexity: the main ideas and some research frontiers. In *Proceedings 12th ISAAC 2001*, Springer-Verlag LNCS 2223, pp. 291–307, 2001.
- [23] M. R. Fellows. Parameterized complexity: the main ideas and connections to practical computing. In *Experimental Algorithmics*, Springer-Verlag LNCS 2547, pp. 51–77. Springer, 2002.

- [24] M. R. Fellows. Blow-ups, win/win's, and crown rules: some new directions in FPT. To appear in *Proceedings 29th WG 2003*, Springer-Verlag LNCS, 2003.
- [25] F. V. Fomin and D. T. Thilikos. Dominating sets in planar graphs: branch-width and exponential speed-up. In *Proceedings 14th ACM-SIAM SODA*, pp. 168–177, 2003.
- [26] F. V. Fomin and D. T. Thilikos. Dominating sets and local treewidth. In *Proceedings 11th ESA 2003*, Springer-Verlag LNCS 2832, 2003.
- [27] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness*. Freeman, 1979.
- [28] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater. *Fundamentals of Domination in Graphs*. Monographs and Textbooks in Pure and Applied Mathematics Vol. 208, Marcel Dekker, 1998.
- [29] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater. *Domination in Graphs*. Monographs and Textbooks in Pure and Applied Mathematics Vol. 209, Marcel Dekker, 1998.
- [30] T. W. Haynes, S. M. Hedetniemi, S. T. Hedetniemi, and M. A. Henning. Domination in graphs applied to electric power networks. *SIAM J. Discrete Math.*, **15(4)**:519–529, 2002.
- [31] D. S. Johnson. Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, **9**: 256–278, 1974.
- [32] I. A. Kanj and L. Perkovic. Improved parameterized algorithms for planar dominating set. In *Proceedings 27th MFCS 2002*, Springer-Verlag LNCS 2420, pp. 399–410, 2002.
- [33] S. Khuller. Algorithms column: the vertex cover problem. *ACM SIGACT News*, **33(2)**:31–33, 2002.
- [34] K. Mehlhorn and S. Näher. *LEDA: A Platform of Combinatorial and Geometric Computing*. Cambridge University Press, 1999.
- [35] G. L. Nemhauser and L. E. Trotter. Vertex packing: structural properties and algorithms. *Mathematical Programming*, **8**:232–248, 1975.
- [36] R. Niedermeier and P. Rossmanith. On efficient fixed-parameter algorithms for Weighted Vertex Cover. *Journal of Algorithms*, **47(2)**:63–77, 2003.
- [37] F. S. Roberts. *Graph Theory and Its Applications to Problems of Society*. SIAM Press 1978. Third printing 1993 by Odyssey Press.
- [38] L. A. Sanchis. Experimental analysis of heuristic algorithms for the Dominating Set problem. *Algorithmica*, **33(1)**:3–18, 2002.

- [39] J. A. Telle. Complexity of domination-type problems in graphs. *Nordic J. Comput.* **1**:157–171, 1994.
- [40] P.-J. Wan, K. M. Alzoubi, and O. Frieder. A simple heuristic for Minimum Connected Dominating Set in graphs. *International Journal of Foundations of Computer Science*, **14(2)**:323–333, 2003.
- [41] K. Weihe. Covering trains by stations or the power of data reduction. In *Proceedings 1st ALEX'98*, pp. 1–8, 1998.
<http://rtm.science.unitn.it/alex98/proceedings.html>.
- [42] K. Weihe. On the differences between “practical” and “applied” (invited paper). In *Proceedings WAE 2000*, Springer-Verlag LNCS 1982, pp. 1–10, 2001.