MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **58** (2007) 445-460

ISSN 0340 - 6253

A knowledge-based system for graph theory, demonstrated by partial proofs for graph-colouring problems

Dieter Gernert
Technische Universität München
Arcisstr. 21, D-80333 München (Germany)
Email: t4141ax@mail.lrz-muenchen.de

Landon Rabern
University of California at Santa Barbara
Santa Barbara CA 93106 (USA)
Email: landonr@math.ucsb.edu

(Received June 21, 2006)

Abstract: *The software system KBGRAPH, which supports graph theoretical proofs and the analysis of graph classes, is presented by developing partial proofs for two graph colouring problems. It is shown that Reed's Conjecture, which concerns an upper bound to the chromatic number, holds for some special classes of graphs; future approaches are briefly outlined. Another strengthening of Brooks' well-known upper bound is sketched. Details about the internal derivation strategies of the program and tools offered to the users are presented, as far as needed for an understanding of the subsequent sketch of a problem-solving process. For readers particularly interested in the software system further hints on its implementation, technical data, and the availability of the program are compiled in the last chapter.*

1. A First Look at Some Open Problems

In the spirit of the conference CSD III, with its focus on the computer-supported generation of hypotheses and computer-assisted proofs in graph theory, we want to let the reader participate in a step-by-step search for new restrictive conditions, aiming at a proof of Reed's Conjecture. Yet only partial proofs have been found. In the course of this stepwise search, also new relations between graph invariants have been discovered, which are valid for all graphs and hence of interest independently.

Reed's Conjecture is an extension of a well-known upper bound for the chromatic number χ (G) (Brooks 1941):

Conjecture (Reed): For any graph G

$$\chi(G) \le \lceil(\Delta(G) + \omega(G) + 1)/2\rceil \qquad (1)$$

where $\Delta(G)$ denotes the maximum degree, and $\omega(G)$ is the clique number of G.

The chromatic number χ is a graph invariant which is connected with a great variety of other invariants, such that sharper bounds for χ may lead to an improved knowledge about other variables; hence χ may also become relevant in mathematical chemistry. In the graph theory textbook by Merris (2001), where most of the examples and applications are taken from chemistry, just the second chapter (after the introduction) is dedicated to graph colourings. Some graph invariants with relevance to mathematical chemistry will be addressed in a proper place (Section 3).

The following sketch will show how structural knowledge on the one hand, and inequalities stored in a special software, on the other hand, can be used to proceed towards partial proofs. The system KBGRAPH, which is designed mainly for the support of graph theoretical proofs, will frequently be "in the background"; anyway the following text is written such that no prior knowledge about that system is needed. Therefore only a necessary short overview is given in the beginning (Section 3), whereas more details for interested readers are placed in a separate chapter (Section 5).

As a by-product, some new inequalities, partially connected with Reed's Conjecture, will be listed separately (Section 4.5), particularly some sufficient conditions for a strengthening of Brooks' inequality such that $\chi \le \Delta$ is replaced by $\chi \le \Delta - 1$.

2. Notation

Here "graph colouring" always refers to the vertex-coloration of graphs. All graphs considered are simple (finite, undirected, without loops and multiple edges). As usual, G is a graph with p vertices and q edges, and G^c is its complement. Special graph classes are the complete graphs K_m, the cycles C_m, and the claw $K_{1,3}$. Frequently occurring variables are the minimum degree δ and the maximum degree Δ, the clique number ω (number of vertices of the greatest clique), the independence number β_0 and the vertex-cover number α_0, the vertex connectivity κ, and the edge connectivity κ_1. We will also write χ instead of $\chi(G)$, etc., if everything is clear.

3. Overview of the System KBGRAPH

The knowledge-based system KBGRAPH pursues two purposes:

- analysis of a given class of graphs, and
- support to proofs of graph theoretical hypotheses.

At the moment, the knowledge base consists of more than 1500 entries. Each such entry has the form of a known property of a graph invariant (e.g., $\omega \ge 2$) or of a relation between graph invariants, which may be unconditional (e.g., $\chi \ge \omega$) or conditional (IF … THEN …). Integer, real, and Boolean variables are permitted, as well as logical connectors (AND, OR, NOT). Each entry (except for some trivial ones) is equipped with a reference.

About 50 graph invariants are implemented, including all which are mentioned in this paper. Graph invariants with particular relevance to chemistry are the greatest, second-greatest, and least adjacency eigenvalue, the greatest and the second-smallest Laplacian eigenvalue, the energy of a graph (the sum of the absolute values of all adjacency eigenvalues), and the sum of all squared degrees (also called the first Zagreb index). Material on the Wiener index and some other variables has been accumulated in the form of paper, but not yet entered into the knowledge base.

A problem description consists of a finite list of "user-defined restrictions". These have the same shape as the knowledge-base entries: conditional or unconditional equations or inequalities; in practice, the unconditional statements are more frequent. Any property of the considered class of graphs which is known in before can be entered here as a user-defined restriction (see the example in Section 4.2). In those cases where an attempt of a mathematical proof is made, a problem description lists known properties of a hypothetical counterexample.

In the very onset of the evaluation process, just after reading the user-defined restrictions, an internal duplicate of the knowledge base is generated, and this is confronted with the user-defined restrictions. The main evaluation process works in the usual mathematical style (forward chaining): known numerical and Boolean values are inserted, and the formulas are simplified. In this way, the temporary duplicate of the knowledge base is permanently updated. As soon as the IF-part of a conditional statement is found to be true, this IF-part is deleted and the THEN-part remains as an unconditional statement. If a THEN-part turns out as false, the negation of the IF-part is kept as a true statement. The transitivity of equality and of inequality relations is taken into account (e.g., $\kappa \le \kappa_1$ and $\kappa_1 \le \delta$ implies $\kappa \le \delta$). A special table is set up and permanently updated which stores the currently best known numerical values for the lower and upper bounds of the numerical variables.

Within the general frame of forward chaining, some specific techniques applied within the inference process can be sketched here (only in a selection):

- rounding in the case of integer variables: e.g., $\chi < 9/2$ is replaced by $\chi \le 4$;
- deletion of formulas which are inferior to other entries in the set of transformed formulas;
- conclusions derived from the monotonicity of arithmetic functions: If, e.g., $y = f(x)$ is a monotonically increasing function for $a \le x \le b$ $(a < b)$, then it can be derived that $f(a) < f(b)$. (Monotonicity of a function can be recognized for linear or quadratic expressions and for functions of the type $y = c*\log(x) + d$.)

(Further special methods of evaluation are described in Section 5.)

When an inference run has ended, then generally concrete values for some variables have been identified, and improved bounds to some numerical variables have been found. These values and bounds are displayed to the user. Another part of the intermediate results consists of those knowledge-base entries which were altered by the evaluation runs. These formulas can optionally be displayed on a screen, either completely or in a selective manner by use of the retrieval function. After any evaluation run, the user may enter further knowledge – possibly triggered by studying the recent results – and restart the evaluation. This may be repeated recursively as long as some improvement will be found by the system.

If, in the case of an attempted proof, a contradiction is found, this means that the underlying class of graphs – that is the class of hypothetical counterexamples - is empty, or, equivalently, that the hypothesis has been proved. This is signalled to the user, together with data about the formulas which led to that contradiction.

KBGRAPH is consequently organized for interactive working. After the end of an evaluation run the user has the chance

- to enter additional knowledge in the same style as the initial user-defined restrictions
- to edit a single formula (e.g., to simplify an arithmetic expression by hand)
- to enter the knowledge that some If-part is true or some THEN-part is false
- to tentatively insert numerical values for a numerical variable (in the case of a contradiction it can be possible to lift a lower or to reduce an upper bound).

For each result, a "derivation tree" can be displayed, which shows how the result has been derived. The numbers of the formulas lead to the references from which formulas of the original knowledge base had been taken.

Some advanced evaluation techniques are implemented, too, e.g., *working with case distinctions.* For Boolean variables, the two alternatives can analysed separately (e.g., *regular / not regular*); in the case of a numerical variable, its domain is decomposed into partial intervals (see an example in Section 4.2). There can be an identical improvement for the different cases, which had been detected in quite distinct ways of derivation. (For further advanced evaluation techniques see Sections 4.2 and 5.2.)

A characteristic phenomenon appearing in evaluation processes can be dubbed *knowledge propagation*: improved knowledge about one variable is likely to advance the knowledge about some other variables. Inspection of the ways in which surprising results come up suggests the term *"crossword-puzzle phenomenon".* When a crossword puzzle is solved, a single new finding can trigger a "chain reaction" of further new findings, such that finally entries for distant places will be found. Hence any raising of a lower or shrinking of an upper bound can be considered a chance for further progress. Furthermore, some conditional formulas will be activated as soon as a bound in a condition will be reached.

The system supplies improved knowledge about the considered class of graphs, in particular: exact values for some graph invariants, sharper bounds for most of the other variables, and restrictions in the form of equations or inequalities to be fulfilled by graph invariants. If no proof is derived (which is the regular case for long-standing graph theoretical conjectures), then the new knowledge about properties of a counterexample may simplify the remaining task.

4. Reed's Conjecture

4.1 Problem Statement

Brooks (1941) proved that for all connected graphs the inequality

$$\chi \leq \Delta + 1 \tag{2}$$

holds, with equality if and only G is a complete graph or an odd cycle (here only the case of connected graphs with $\chi \leq \Delta$ is of interest). This was strengthened by Reed (1998, 1999), whose conjecture was quoted above as (1). Neither a proof nor a counterexample are known. This conjecture is trivial for $\omega = \Delta$ and for $\omega = \Delta + 1$. Reed (1998) gave a proof for graphs with a maximum degree $\Delta = p - 1$. A proof for all line-graphs was presented at a conference in Berlin (June 2005, see King et al. (2005)); this proof stood in the context of a harder claim on multi-edge graphs. A quick proof for line-graphs - restricted to simple graphs - will be obtained below as a by-product (Section 4.2, step 2). Some contributions in two recent papers (Rabern 2006a, b) will be compiled below (Section 4.3).

4.2 A First Attempt with Reed's Conjecture

Unless otherwise stated, G is to denote a counterexample to Reed's Conjecture. It is our goal to find more and sharper constraints which the class of counterexamples will have to fulfil.

Step 1: We can restrict our study to colour-critical graphs with chromatic number χ (χ-critical graphs). Every graph with chromatic number χ contains a χ-critical subgraph with the same number of vertices. If such a graph obeys (1), then in any other graph generated from it by inserting some edges, Δ and ω can only remain constant or increase, such that (1) will remain valid. So we can make use of the known properties of colour-critical graphs. (The property "colour-critical" cannot be found automatically, but it is implemented in the system as a Boolean variable and it will be used if stated by the user.)

Step 2: By inserting $\omega = \chi$ and $\omega = \chi - 1$ into (1), it turns out that (1) is fulfilled for these values. Hence G has to satisfy the constraint

$$\chi \geq \omega + 2 \tag{3}$$

Some consequences of this property are stored in the system. Furthermore, it is known from the literature (Kierstead and Schmerl 1986) that graphs obeying (3) must contain $K_{1,3}$ and/or K_5-e as an induced subgraph. These graphs are forbidden induced subgraphs for line-graphs, and so it is quickly proved that (1) holds for line-graphs.

Here the restriction (3) was found by the user. In principle, it would be possible to start a program run without entering (3) as a user-defined restriction – the program would be able to exclude $\omega = \chi$ and $\omega = \chi - 1$ in later phases. But by doing so, the program run would become longer, and the intended demonstration would become rather clumsy. Furthermore, it should also be shown that the user's additional prior knowledge can be entered here.

Step 3: Next we will check whether $\chi = \omega + 2$, $\omega = \chi - 2$ is possible. Since counterexamples are studied we have

$$\chi > \lceil (\Delta + \omega + 1)/2 \rceil$$

or equivalently

$\chi > (\Delta + \omega + 1 + \varepsilon)/2,$
$\chi > \Delta - 1 + \varepsilon$

with $\varepsilon = 1$ if $\Delta + \omega$ is even, and $\varepsilon = 0$ if $\Delta + \omega$ is odd. The case $\varepsilon = 1$ must be excluded since here $\chi \leq \Delta$. Only the case $\varepsilon = 0$ remains. Then $\chi = \Delta$, and $\chi = \omega + 2$ implies $\chi \equiv \omega \pmod 2$, $\Delta \equiv \omega \pmod 2$, $\Delta + \omega \equiv 0 \pmod 2$, $\varepsilon = 1$, contrary to $\varepsilon = 0$. Hence $\chi = \omega + 2$ is excluded, and with (3) we obtain

$$\chi \geq \omega + 3 \tag{4}$$

This derivation cannot be accomplished by the system.

Step 4: As a next step we can compile the user-defined restrictions:

R1: $\chi > \lceil(\Delta + \omega + 1)/2\rceil$
R2: $\Delta \leq p - 2$
R3: colour-critical
R4: $\chi \geq \omega + 3$

Here R1 is the negation of (1) since we are looking for a counterexample. R2 is a consequence of Reed's additional restriction as cited above. R3 was explained before (Step 1), and R4 goes back to Step 3.

Step 5: With these user-defined restrictions a first evaluation run is started. Among the results only two points are worth reporting: a counterexample G is not completely multipartite, and it has $p \geq 11$. The latter finding is mainly due to a theorem by Nenov (1998): here $\chi \geq 5$, and for $\omega \leq 3$ and $p \leq 10$ it would follow that $\chi \leq 4$; for omega $\omega \geq 4$, with $\chi \geq 7$, the derivation is different.

Step 6: After the end of the first standard evaluation run, the special evaluation technique "working with case distinctions" is activated. We consider the complete case distinction $\{\omega = 2,\ 3,\ 4,\ \geq 5\}$, which means that the program will consecutively (but independently) handle the four cases

$\omega = 2, \quad \omega = 3, \quad \omega = 4, \quad \omega \geq 5.$

A selection of the results is given by the table:

$\omega =$	2	3	4	≥ 5
$p \geq$	22	12	13	15
$q \geq$	47	34	43	58
$\chi \geq$	5	6	7	8
$\gamma \geq$	2	1	2	2

where γ is the orientable genus. First, we notice a little improvement from $p \geq 11$ to $p \geq 12$. The lower bound $p \geq 22$ for $\omega = 2$ can be traced back to a result by Jensen and Royle (1995): a graph with $\omega = 2$ and $\chi = 5$ has $p \geq 22$ (this lower bound also holds for $\chi \geq 6$).For the lower bounds to q, there is quite a lot of inequalities in the knowledge base, some of which also use parameters like ω and Δ; one of the most efficient lower bound for q in the case of colour-critical-graphs was found by Kostochka et al. (1999, see Section 4.5, nr. 2). The lower bounds for χ follow from (4).

In three of the four cases we have $\gamma \geq 2$. Before we proceed to the case with $\gamma \geq 1$, we can study the functioning of the program module "working with case distinctions". The same result, $\gamma \geq 2$, was derived in three different ways for the three cases (at the same time, it can be noticed here that after a program run the derivation of each result can be displayed). In this particular case, we can identify the underlying knowledge-base entries:

- Case 1, $\omega = 2$:
  IF $\gamma \leq 1$ and $\omega = 2$ THEN $\chi \leq 4$ (Kronk 1969)
- Case 3: $\omega = 4$:
  IF $\chi \geq 7$ and $\omega \leq 6$ THEN $\gamma \geq 2$ (Thomassen 1994)
- Case 4: $\omega \geq 5$:
  IF $\gamma \leq 1$ THEN $\chi \leq 7$ (Kronk 1969)

Step 7: In view of the preliminary lower bounds for γ, the user may decide to handle the troublemaker – Case 2 with $\gamma \geq 1$ – separately. To this purpose, a new program run is started for Case 2 with the "hypothesis" $\gamma = 1$ as a new user-defined restriction. This program run uses all known results for this case and all restrictions defined earlier, particularly $\omega = 3$ and $\chi \geq 6$. Mainly on the basis of a formula by Dirac (1952), the program supplies $\chi = \delta = \Delta = 6$, such that G would be regular. But according to Gould (1988, p. 247), a colour-critical graph with $\delta \geq 3$, $\delta = \chi$ cannot be regular. This contradiction, which is displayed to the user, excludes $\gamma = 1$, and so $\gamma \geq 2$ has been proved for this case.

Just to illustrate the flexibility of the system, we show an alternative proof for $\gamma \geq 2$ in Case 2: a "semi-automatic", computer-assisted proof, which starts from the known facts $\omega = 3$ and $\chi \geq 6$. The retrieval function is activated, and as a response to the query "χ and γ", about 20 formulas containing χ and γ are displayed on the screen. Stimulated by a formula due to Thomassen (1994) the user can look up the original printed version. According to this source, most of the graphs with $\gamma = 1$ have $\chi \leq 5$, and hence can be ignored here. Two exceptional graphs have $\omega \geq 4$, contrary to $\omega = 3$. For the third of Thomassen's exceptional cases, one can combine the fact that here $p \geq 12$ with findings by Albertson and Hutchinson (1980), with the consequence that also this last exceptional graph can be omitted. So we derived that G has $\gamma \geq 2$, or, in other words, that (1) holds for planar and for toroidal graphs.

4.3 Contributions from the Theory of Graph Associations

The following inequalities, valid for all graphs, are taken from two papers recently published or just under the press (Rabern 2006a, b). Through the concept of graph associations, a theorem is found which permits us to derive new bounds for χ by choosing special types of induced subgraphs.

Definition: Given a graph G and non-adjacent vertices a and b, we write G/[a,b] for the graph obtained from G by associating (i.e. identifying) a and b into a single vertex [a,b] and discarding multiple edges.

Theorem 1: Let G be a graph. Then, for any induced subgraph H of G

$$\chi(G) \leq \chi(H) + (p(G) + \omega(G) - p(H) - 1)/2 \tag{5}$$

There are two immediate applications. If G is connected, and if the subgraph H is identified with a longest induced path P_m of G ($m \geq 3$, such that the diameter $d(P_m) = d(G) = m - 1 \geq 2$), then (5) leads to

$$\chi(G) \leq (p(G) + \omega(G) - d(G) + 2)/2 \tag{6}$$

Next, suppose that G has $g \geq 5$ (where the girth g is the length of the shortest cycle), and take for H a subgraph induced by a shortest cycle together with its neighbourhood, then

$$\chi(G) \leq (p(G) - g(G)(\delta(G) - 1) + 7)/2 \tag{7}$$

Theorem 2: Let G be a graph. Then

$$\chi(G) \leq (p(G) + \omega(G) - \beta_0(G) + 1)/2 \tag{8}$$

or equivalently (with Gallai's relation $\alpha_0 + \beta_0 = p$)

$$\chi(G) \leq (\omega(G) + \alpha_0(G) + 1)/2 \tag{9}$$

For triangle-free graphs it follows that

$$\chi(G) \leq (p(G) - \Delta(G) + 3)/2 \tag{10}$$

The following three inequalities are related to Reed's Conjecture. It was found that (1) holds for *decomposable graphs*, that is for graphs G with a disconnected complement G^c, such that G can be written as a direct sum $G = A + B$ ($A + B$ means that each vertex of A is connected with each vertex of B). If G is a counterexample to (1), then G^c has a perfect matching if p is even; for odd p, G^c is nearly matching-covered, that is, G^c-v has a perfect matching for any vertex v (see also Stehlík (2003)); furthermore G^c is bridgeless, that is, $\kappa_1(G^c) \geq 2$.

Any counterexample to (1) must satisfy the inequalities

$$\chi(G) \leq \lceil p(G)/2 \rceil \tag{11}$$
$$\Delta(G) \leq p(G) - \mathrm{sqrt}(p(G) + 2\beta_0(G) + 1) \tag{12}$$
$$\beta_0(G) \geq 3 \tag{13}$$

4.4 Example for an Advanced Evaluation Technique

As remarked before, a proof for Reed's Conjecture has not yet been found. Some experimenting with the program (and material from literature) suggest that there are two graph classes for which a solution may seem relatively easy: the *triangle-free graphs* and the *claw-free graphs*. The following episode is to show – based on an example – one of the advanced features of the program, which the user can apply after the usual inference runs.

Triangle-free graphs are characterized by $\omega = 2$. After an ordinary program run (for counterexamples to Reed's Conjecture), including the new constraint $\omega = 2$, a run with a case distinction (cf. Section 4.2, step 6) followed. The case distinction $\{g = 4, \ g \geq 5\}$ led to the result that $\gamma \geq 2$ for $g = 4$, and $\gamma \geq 3$ for $g \geq 5$. This suggests a test whether $\gamma \geq 3$ could be proved for $g = 4$, too. The output of a new program run with the recent constraints $g = 4$ and $\gamma = 2$ consists of fixed values for nine numerical variables (e.g., $\chi = \Delta = 5$), whereas for all other numerical variables both lower and upper bounds are supplied.

So this is an ideal candidate for a program function called "automatic insertion": for each integer variable which is constrained from both sides, all admissible values are inserted into the formulas of the knowledge base, and a contradiction close to a bound leads to an increase of a lower or a decrease of an upper bound (an extension to intervals bounded at one side and to real variables can only be mentioned here, cf. Section 5.2). In the concrete case, e.g., the interval $22 \leq p \leq 86$ was replaced by $22 \leq p \leq 56$, and $47 \leq q \leq 176$ was converted into $47 \leq q \leq 116$; in a similar way, the inclusions for most of the other integer variables were strengthened.

Other advanced techniques of evaluation were applied to Reed's Conjecture, too, but in this special case they did not lead to significant progress. Therefore these techniques will be handled in a general form in Section 5.2; a summary of partial results is to follow in Section 4.6.

4.5 Miscellaneous Inequalities

Some related inequalities, vastly scattered in literature, should be registered here (with a short derivation or reference). The following formulas were found by the retrieval function of KBGRAPH upon the query "χ and Δ". Some of them have immediate consequences for Reed's Conjecture. (The references are supplied by the system; a selection and some editing for the sake of easy reading were required.)

1. Brooks' well-known result, that all connected graphs (except for complete graphs and odd cycles) satisfy $\chi \leq \Delta$, suggests to ask for conditions under which the stronger property

$$\chi \leq \Delta - 1 \tag{14}$$

will hold. Some sufficient conditions are:

a.) If $\omega = 2$ then $\chi \leq 2(\Delta + 2)/3 + 1$
(Stacho 2001)
$\rightarrow$ $\omega = 2$ and $\Delta \geq 8$ implies (14).

b.) If $\omega \leq 3$ then $\chi \leq 3(\Delta + 2)/4$

(Borodin et al. 1977)
$\rightarrow$ $\omega \leq 3$ and $\Delta \geq 7$ implies (14).
c.) If $\Delta \geq 7$ and $\omega \leq (\Delta - 1)/2$ then $\chi \leq \Delta - 1$
(Borodin et al. 1977)
d.) If $\chi \geq \omega + 1$ and $\Delta > (p + 1)/2$ then $\chi \leq \Delta - 1$
If $\chi \geq \omega + 1$ and $\Delta \geq 9$ and $\Delta > p/2$ then $\chi \leq \Delta - 1$
(Beutelspacher and Hering 1983)
e.) If G has no C_4 (induced subgraph or not) then $\chi \leq 2\Delta/3 + 2$
(Catlin 1978)
$\rightarrow$ If G has no C_4 then $\Delta \geq 7$ implies (14).
f.) If G is colour-critical, and $\kappa = 2$ with a cutset {u, v}, then u and v are non-adjacent, and $\Delta \geq (3\chi - 5)/2$ (Gould 1988, p. 227)
$\rightarrow$ Under these conditions $\Delta \geq 6$ implies (14).

2. For colour-critical graphs, given χ and p, a good lower bound for q is required. At present, the best such bound is given by Kostochka et al. (1999), under the reservation that two classes of exceptional graphs are excluded; this restriction can be by expressed by the three "or"-connected properties:
If G is colour-critical and $4 \leq \chi \leq p - 2$ and ($2\chi \neq p + 1$ or $\beta_0 \geq 3$ or $\omega < (p - 1)/2$)
then $q \geq p(\chi - 1)/2 + \chi - 3$

3. If G contains neither C_4 nor $2K_2$ as induced subgraphs, then $\chi(G) + \chi(G^c) \geq p(G)$ and $\chi(G) \leq \omega(G) + 1$ (Blázsik et al. 1993). So, due to (3), also for this special class of graphs (1) holds.

4.6 Summary of Partial Results

Reed's Conjecture was proved for the following classes of graphs:

- line-graphs
- graphs with $\chi \leq \omega + 2$
- completely multipartite graphs
- planar and toroidal graphs
- decomposable graphs
- $\{C_4, 2K_2\}$-free graphs.

In any counterexample, variables have to satisfy the following lower bounds (a small selection): $p \geq 12$, $q \geq 34$, $\chi \geq 5$, $\delta \geq 4$, $\Delta \geq 5$, $\beta_0 \geq 3$, $\alpha_0 \geq 8$, $\gamma \geq 2$. Necessary properties of the complement G^c are compiled in Section 4.3.

5. Further Details on the System KBGRAPH

5.1 Starting Point and General Properties

The project KBGRAPH was started in 1985, stimulated by the appearance of a compilation of relations between graph invariants (Brigham and Dutton 1985). This list with 262 entries –an offspring of the project INGRID - formed the core of the knowledge base in the first version of KBGRAPH. A publication on INGRID (INteractive GRaph Invariant Delimiter) had appeared earlier (Dutton and Brigham 1983); a comprehensive report was finished in 1986, but printed later (Dutton, Brigham and Gomez 1989). Apparently, the project INGRID was no more continued after the publication of a supplement to the compilation quoted before (nr. 263 – 458; Brigham and Dutton 1991).

Although KBGRAPH owes much to the publications just cited, it has been independently developed further. Whereas forward chaining (Section 3) has been maintained as the central evaluation strategy, KBGRAPH is now characterized by a quantitative increase (number of graph invariants and size of the knowledge bases) and by a series of novel features, mainly related to

- the user interface and the options for flexible post-processing,
- the advanced evaluation techniques (Section 5.2),
- the options for an external control of the inference process (Section 5.3).

At the moment, 51 graph invariants are implemented. The three knowledge bases include about 2000 entries: about 1500 in the "main knowledge base", and the rest in two "auxiliary knowledge bases" required for one of the special evaluation techniques (Section 5.2). According to individual requirements, graph invariants can be newly defined, cancelled, or renamed. The knowledge-base is permanently updated: adding, deleting or altering of entries is possible. From time to time a single entry is replaced by a stronger version.

Based upon forward chaining as the central inference method (thus following the example of Brigham and Dutton), the inference mechanism was programmed *ad hoc,* to adapt to the specific requirements of working with formulas (no foreign software was used). Options for an external control of the inference will be outlined in Section 5.3.

After the end of an inference run, the user can enter additional knowledge and start the inference process again, or apply one or the other of the "advanced evaluation techniques" (Section 5.2); all this can be done repeatedly, as long as some progress is expected. After the end of each inference run, it is possible to display a derivation tree for each single result, and, in the case where more than one way of derivation led to the same single result, this fact is disclosed to the user, too. So all findings can be checked and rewritten in the usual mathematical style.

5.2 Advanced Evaluation Techniques

After an ordinary program run the user may decide to use one of the following special techniques, which are all optional:

- working with case distinctions
- automatic insertion of values
- editing of a formula
- transition to a "related graph".

Working with case distinctions was already explained and illustrated in Section 4.2 (step 6). The user is free to define a decomposition of the domain of a variable (up to 9 segments). No closed interval is required – a decomposition can have the shapes, e.g., $\omega = \{2, 3, \geq 4\}$ or $\{3 \leq \Delta \leq 6,\ \Delta \geq 7\}$. The decomposition into sub-classes, sub-sub-classes, … is supported by the system up to 4 hierarchy levels; within the same hierarchy level up to 9 descendants of the same direct ancestor are permitted. If the same improvement is achieved for all subcases of the same case, then this new knowledge can be "reached upward" to the next common ancestor. The idea behind this – supported by experience – is the chance that the same improvement can be derived in different ways within the different subcases. Optionally, the system can make proposals for plausible case distinctions.

The *automatic insertion of values* was exemplified in Section 4.4. It should be supplemented here that no closed interval is required. If an integer variable is bounded only from one side, then the tentative insertion of numerical values starts at that bound, and continues as long as the formula just considered leads to a contradiction and thus permits to narrow that bound. For the case of very large intervals and/or real variables, heuristic procedures exist, which supply preliminary data to the user, who has to decide whether a proposed problem reduction seems plausible.

Editing is possible for each of the formulas which were transformed by the inference run. The user can

- simplify an arithmetic expression "by hand",
- insert numerical or Boolean values for a variable,
- delete an IF-part if it is considered true,
- replace a THEN-part by "false",
- delete a formula (e.g., if it is recognized that an IF-part cannot be satisfied, or that an inequality is inferior to another one – the latter point is supported by the system).

Transition to a "related graph": Some successful proofs in graph theory show that a transition from the given class of graphs to *another class* - called "related graphs" for short - may be advantageous. Such a transition can be defined by any unique unary graph transformation. There are formulas which connect variables of the original graphs with variables of the "related graphs" - by the example of a transition to the complementary graph: theorems of the Nordhaus-Gaddum type or formulas like $\omega(G) = \beta_0(G^c)$. After an ordinary inference run, which yields new information on the original class of graphs, the user may switch over to a class of "related graphs" in order to start an inference process with respect to that second class. Then the new knowledge about the second class of graphs can be automatically transferred back to the original class of graph. The transitions to *complementary graphs* and to *line-graphs* are implemented in the system. The required "interconnection knowledge bases" exist; these are the two "auxiliary knowledge bases" mentioned before. The user is free to define further types of derived graphs; in this case, of course, a corresponding interconnection knowledge base must be set up.

5.3 Options for an External Control of the Inference Process

The essential options for an external influence on the inference process are:

- masking
- ranking of the variables

- working with or without a derivation tree
- partitioning of the knowledge base.

Masking: Each graph invariant can be "masked", that is, it will be treated as inexistent during the same session. This tool is mainly used if the user is sure that a certain variable will not contribute to the solution. Also each single statement can be masked: so it is possible to make an inference run "with or without use of the 4CC".

Ranking of the variables: A ranking, i.e. a linear order, of all graph invariants is defined. In the case where an equality between two numerical variables is derived in the inference process, it is ruled by this ranking whether x will be substituted for y or vice versa. This ranking also has an influence upon the order within output lists. The user can alter the ranking individually and store the new ranking for the future.

Working with or without a derivation tree: The user can decide whether or not a derivation tree is to be built up during an inference run. The derivation tree is required if the user later on wants to get information about the way how a certain result had been derived. Working without the derivation tree will cut the program runtime.

Partitioning of the knowledge base: In view of the extended knowledge bases, strategies are recommendable to speed up the inference process. The main knowledge base is partitioned in the following way. Each of its entries is assigned to one of three subsets whose members may be named "very important", "important", and "less important". In the beginning, only the first-class statements are used; in later inference runs also those in the second class will join, until finally all entries in the knowledge base will be active. In this way, some useful intermediate results can be achieved already in earlier inference runs, with the consequence that many expressions can be simplified rather soon. Four different strategies for a partitioning of the knowledge base were empirically tested. It was found that in most cases this technique leads to a considerable reduction of computing time. The user may choose among these four strategies – anyway, one of them is predefined as a standard (according to the empirical results). For the details only a reference can be given (Gernert 1993).

5.4 Technical Details, References, Availability

The implementation of the system started in 1985 and continued until 2000. Since 2000, no further revision of the program was possible, but the knowledge bases are permanently updated. Due to side conditions about the year 1985 (students' knowledge and equipment, also some administrative rules), the system was programmed in PASCAL (in the final stage about 30000 lines of code) and on the basis of MS-DOS; the menus are in German. A transfer to modern computers is possible, and has already been successfully performed.

Further information, e.g., on examples of application, criteria for the selection of graph invariants and formulas, parameter dependence of required computer runtime, and strategies to speed up computer runs with large knowledge bases, can be found in two papers listed below (Gernert 1989, 1999, with many references). The executable program is available for everybody (email: t4141ax@mail.lrz-muenchen.de), as well as the published and unpublished expertise acquired in many years of practical work with the system.

6. Short Outlook

Future research on Reed's Conjecture may start from the abundant literature on colour-critical graphs, of which only a small proportion was used until now. Another promising approach can be based upon the complements of possible counterexamples – some structural properties of these complementary graphs are compiled here.

The practical use of the system KBGRAPH is going on. The knowledge bases are permanently updated; but nevertheless the system should by programmed totally from the beginning, free from restrictions imposed by earlier hardware, based upon a modern programming language and operating system, and exploiting the expertise accumulated over the years with respect to design, updating, and practical work.

Note added in proof: Meanwhile, by methods as shown above (Sections 4.2 and 4.4), the open case p = 12 could be settled. Reed's Conjecture is valid for all graphs with at most 12 vertices.

References

Albertson, M.O. and Hutchinson, J.P. (1980), On six-chromatic toroidal graphs. Proceedings of the London Mathematical Society (3) **41**, 533-556.

Beutelspacher, A. and Hering, P.-R. (1983), Minimal graphs for which the chromatic number equals the maximal degree. Ars Combinatoria **18**, 201-216.

Blázsik, Z., Hujter, M., Pluhár, A., and Tuza, Zs. (1993), Graphs with no induced C_4 and $2K_2$. Discrete Mathematics **115**, 51-55.

Borodin, O.V. and Kostochka, A.V. (1977), An upper bound of the graph's chromatic number, depending on the graph's degree and density. Journal of Combinatorial Theory **B 23**, 247-250.

Brigham, R.C. and Dutton, R.D. (1985), A compilation of relations between graph invariants. Networks **15**, 73-107.

Brigham, R.C. and Dutton, R.D. (1991), A compilation of relations between graph invariants – supplement I. Networks **21**, 412-455.

Brooks, R.L. (1941), On colouring the nodes of a network. Proceedings of the Cambridge Philosophical Society **37**, 194-197.

Catlin, P.A. (1978), Another bound on the chromatic number of a graph. Discrete Mathematics **24**, 1-6.

Dirac, G.A. (1952), Map-colour theorems. Canadian Journal of Mathematics **4**, 480-490.

Dutton, R.D. and Brigham, R.C. (1983), INGRID: a software tool for extremal graph theory research, Congressus Numerantium **39**, 337-352.

Dutton, R.D., Brigham, R.C. and Gomez, F, (1989). INGRID: a graph invariant manipulator. Journal of Symbolic Computation **7**, 163-177.

Gernert, D. (1989), A knowledge-based system for graph theory. Methods of Operations Research **63**, 457-464.

Gernert, D. (1993), Experimental results on the efficiency of rule-based systems. In: A Karmann et al. (Eds.), Operations Research '92. Physica-Verlag, Heidelberg, pp. 262-264.

Gernert, D. (1999), Cognitive aspects of very large knowledge-based systems. Cognitive Systems **5**, 113-122.

Gould, R. (1988), Graph theory. Benjamin Publ. Comp., Menlo Park CA.

Jensen, T. and Royle, G.F. (1995), Small graphs with chromatic number 5: a computer search. Journal of Graph Theory **19**, 107-116.

Kierstead, H.A. and Schmerl, J.H. (1986), The chromatic number of graphs which neither induce $K_{1,3}$ nor K_5-e. Discrete Mathematics **58** (1986) 253-262.

King, A.D., Reed, B.A. and Vetta, A. (2005), An upper bound for the chromatic number of line graphs. Lecture given at EuroComb 2005 (Berlin, June 2005), DMTCS Proc. AE 2005, 151-156.

Kostochka, A.V. and Stiebitz, M. (1999), Excess in colour-critical graphs. In: Graph theory and combinatorial biology (Proc. Balatonlelle (Hungary) 1996). Bolyai Society Mathematical Studies **7**, 87-99.

Kronk, H.V. (1972), The chromatic number of triangle-free graphs. Lecture Notes in Mathematics **303**, 179-181.

Merris, R. (2001), Graph theory. Wiley, New York.

Nenov, N.D. (1998), On the small graphs with chromatic number 5 without 4-cliques. Discrete Mathematics **188**, 297-298.

Rabern, L. (2006a), On graph associations. SIAM Journal on Discrete Mathematics **20**, 529-535.

Rabern, L. (2006b), A note on Reed's Conjecture, arXiv math.CO/0604499.

Reed, B. (1998), ω, Δ, and χ. Journal of Graph Theory **27**, 177-212.

Reed, B. (1999), A strengthening of Brooks' theorem. Journal of Combinatorial Theory **B 76**, 136-149.

Stacho, L. (2001), New upper bounds for the chromatic number of a graph. Journal of Graph Theory **36**, 117-120.

Stehlík, M. (2003), Critical graphs with connected complements. Journal of Combinatorial Theory **B 89**, 189-194.

Thomassen, C. (1994), Five-coloring graphs on the torus. Journal of Combinatorial Theory **B 62**, 11-33.