Belgian Journal of Operations Research, Statistics and Computer Science Vol. 40 (1-2) 00

Pentomino exclusion and spanning. IP-formulation, valid inequalities and facets

Frank Plastria

Department of Management Informatics

Research Group for Location and Distribution

Vrije Universiteit Brussel

Pleinlaan 2, B 1050 Brussels, Belgium

e-mail: Frank.Plastria@vub.ac.be

January 24, 2001

Abstract

Pentomino exclusion asks to delete a minimum number of cells from a square grid forbidding placement of any pentomino shape within the remaining cells, while pentomino spanning asks for finding a minimum number of different disjoint pentominoes disallowing placement of any additional pentomino. We discuss here IP formulations for each of these two problems of recreational mathematics. Several families of symmetry breaking constraints and valid inequalities are derived, many of which are shown to be facet generating.

1 Introduction

One way of catching the attention of students and motivating them for theoretical developments they otherwise often feel too terse, is to make use of examples from recreational mathematics without other direct practical application. The area of combinatorial puzzles offers a quite rich collection of such examples particularly suited to the illustration of integer programming, and many topics in this field as developed in e.g. [6]. This paper examines two such problems which were recently posed as brain teasers in Optima, the newsletter of the Mathematical Programming Society. The aim is to show how they can illustrate modelling with binary variables, valid inequalities and facets. Therefore the paper does not include computational results.

A pentomino (hexomino) is a connected union of five (six) square cells out of a regular planar square grid. Two pentominoes are equivalent when they have the same shape, i.e. when they can be obtained from each other after applying any plane symmetry. There exist exactly 12 distinct pentomino shapes (see table 1) and a well known puzzle consists in placing all these within a 8×8 board, leaving 4 specific cells free, e.g. the corner ones, or the center ones.

In [2] R.A.Bosch poses the following two lesser known problems as brain teasers:

Pentomino Exclusion A set of monominoes (cells) of the $n \times n$ board excludes the pentominoes when they leave no room for any pentomino. What is the smallest set excluding all pentominoes ? Devise an Integer Programming (IP) formulation for this problem.

Pentomino Spanning A subset of the pentominoes spans a board if its members can be placed on the board so that they disallow any placement of the remaining pentominoes. Devise an IP formulation for finding the smallest set of pentominoes that spans the $n \times n$ board.

For the first problem [2] suggests a zero-one linear programming formulation, further improved by a few additional valid inequalities. In the first part of this paper we extend the study of this formulation in several ways. On the one hand we introduce additional symmetry breaking constraints to limit the feasible space in order to avoid equivalent solutions due to the symmetry inherent in the problem. On the other hand we construct several families of valid inequalities by inspecting subsets of the board which are hexominoes. We then show that many (but not all) of these are facet generating.

The second part of the paper examines the pentomino spanning problem in a similar manner. It is first formulated as a zero-one linear programme, symmetry breaking constraints are then introduced, and valid inequalities are derived.

The last section suggests several possible extensions of the studied problems.

2 The pentomino exclusion problem

2.1 Definition and simple model

The following is a reformulation of the model described in [2].

Consider the set I of cells of the board, and with each cell $i \in I$ we associate a binary variable $x_i \in \{0,1\}$, stating whether a monomino has been placed on cell i ($x_i = 1$) or not ($x_i = 0$) . In the first case we say that cell i is used, in the latter that cell i is free.

A placement P of a pentomino on the board is a subset of the cells of the board with a particular shape : $P \subset I$ is connected and contains 5 cells. The set of all possible pentomino placements is therefore a set $\mathcal{P}$ of subsets of I.

Pentominoes are connected shapes consisting of five squares. Barring symmetry, twelve pentominoes exist, given in Table 1. First a capital letter is given as a (tentative) name, which resembles the actual shape shown in the second column. Next the table indicates (and is sorted on) their number of forms (different symmetric forms the shape can take through some isometry of the square (rotation and/or mirroring), i.e. 8 divided by the shape's number of symmetries), the dimensions of their rectangle-hull (h, w), the number of possible placements of each of their forms separately on a $n \times n$ board $(n - h + 1)(n - w + 1)$, and the total number of placements of the shape (the product of the two previous values). Note that we assume $n \geq 5$, since some shapes do not fit on smaller boards.

A set of monominoes (or a value- setting of the x_i) excludes the placement P if at least one of the cells of P is used. The *pentomino exclusion problem* now asks for the minimal number of monominoes which excludes all pentomino placements.

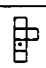
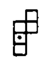
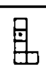
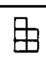
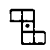
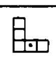
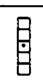
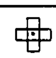
This is directly formulated as an IP as follows :

$$\min \sum_{i \in I} x_i \tag{1}$$

$$\sum_{i \in P} x_i \geq 1 \qquad \forall P \in \mathcal{P} \tag{2}$$

$$x_i \in \{0,1\} \qquad \forall i \in I \tag{3}$$

Table 1: list of all pentominoes

Name	Shape	# forms	box	#plac/form	#placements
R		8	(3,3)	$(n-2)^2$	$8n^2 - 32n + 32$
Y		8	(4,2)	$(n-3)(n-1)$	$8n^2 - 32n + 24$
S		8	(4,2)	$(n-3)(n-1)$	$8n^2 - 32n + 24$
L		8	(4,2)	$(n-3)(n-1)$	$8n^2 - 32n + 24$
B		8	(3,2)	$(n-2)(n-1)$	$8n^2 - 24n + 16$
Z		4	(3,3)	$(n-2)^2$	$4n^2 - 16n + 16$
W		4	(3,3)	$(n-2)^2$	$4n^2 - 16n + 16$
T		4	(3,3)	$(n-2)^2$	$4n^2 - 16n + 16$
V		4	(3,3)	$(n-2)^2$	$4n^2 - 16n + 16$
C		4	(3,2)	$(n-2)(n-1)$	$4n^2 - 12n + 8$
I		2	(5,1)	$(n-4)n$	$2n^2 - 8n$
X		1	(3,3)	$(n-2)^2$	$n^2 - 4n + 4$
Total					$63n^2 - 240n + 196$

which can be viewed as an instance of the general set covering problem.

It follows that this formulation of the pentomino exclusion problem on a $n \times n$ board ($n \geq 5$) has n^2 binary variables and $63n^2 - 240n + 196$ constraints. For $n = 8$ we have 64 binary variables and 2308 constraints. The LP relaxation of this 8×8 board case has an optimal solution given by $x_i = 0.2$ for all $i \in I$, with objective value 12.8, very far from the integer solution with value 24, given in [2], who also report that CPLEX ran out of memory when attempting to solve the IP by standard branch and bound.

The large integrality gap may be seen as one reason for this failure, the other one being the fact that the objective value does not differentiate much between all solutions: only 64 objective values are possible for the $2^{64} \approx 10^{20}$ (unconstrained) solutions. It follows that there are very many solutions yielding a same objective value, a property we will call 'solution degeneracy'.

This information may also be exploited. The knowledge of integrality of the optimal value allows the branch and bound process to fathom on the threshold value of 1 less than the incumbent. Most commercial IP software seem to allow for this improved fathoming test, but not all. E.g. the Excel Solver in its shrink-wrapped version does not, but allows to indicate a relative 'tolerance' only. This may also be exploited in a slightly less efficient way as follows. Evidently a simple checkerboard pattern of alternating used and free cells is a feasible solution with objective value $\lfloor \frac{n^2}{2} \rfloor$ (if n is odd we, of course, prefer to have more free cells than used ones) and is therefore an upper bound to the optimal value — better bounds may be found, e.g. the 2×2 free square pattern given in [2] with (optimal) value 24 for $n = 8$. Any optimal value being an integer it is at least better by a factor of $1/\lfloor \frac{n^2}{2} \rfloor$, a value that may therefore be used as relative tolerance.

Therefore two avenues of improvement should be studied aimed at improving on both levels. We first investigate how to reduce the number of equivalent solutions by symmetry breaking. Then we study how to reduce the integrality gap, i.e. tighten the description of the feasible set by way of additional valid inequalities.

2.2 Symmetry breaking

A square board possesses a symmetry group of 8 elements, e.g. generated by the horizontal, vertical and diagonal flips. Therefore to any configuration there possibly are up to 8 symmetrical configurations which are essentially equivalent. An evident step in reducing the solution degeneracy is to introduce *symmetry breaking* constraints which allow for at least one representant configuration, but rule out as unfeasible all its symmetrical ones.

In [2] the following symmetry breaking constraints are suggested: there should be at least as many chosen cells in the upper (left) halfsquare than in the lower (right) half. Indeed, if this were not the case one (or two) corresponding flip constructs an equivalent solution satisfying these constraints. These constraints are easily written as follows, using momentarily the for standard boards more common (but often cumbersome) row-column notation $I = \{(r, c) \mid 1 \leq r, c \leq n\}$, and $x_{(r,c)}$ is written shortly as x_{rc}.

$$\sum_{r=1}^{\lfloor \frac{n}{2} \rfloor} \sum_{c=1}^{n} x_{rc} \geq \sum_{r=\lceil \frac{n}{2} \rceil}^{n} \sum_{c=1}^{n} x_{rc} \tag{4}$$

$$\sum_{c=1}^{\lfloor \frac{n}{2} \rfloor} \sum_{r=1}^{n} x_{rc} \geq \sum_{c=\lceil \frac{n}{2} \rceil}^{n} \sum_{r=1}^{n} x_{rc} \tag{5}$$

In fact these constraints do not carry out the full symmetry breaking job, and this for two reasons :

- Constraints (4) and (5) only break symmetry for the horizontal and vertical reflexion, which generate only half of the square's symmetry group. To break the full symmetry group, one should add a constraint to break one more (and last) generator of the group, which leaves the upper and left halfboards invariant: the reflexion w.r.t. the second diagonal (NW-SE). So we propose addition of the following constraint which calls for at least as many used cells in the NE triangle than in the SW triangle.

$$\sum_{c=2}^{n} \sum_{r=1}^{c-1} x_{rc} \geq \sum_{c=1}^{n-1} \sum_{r=c+1}^{n} x_{rc} \tag{6}$$

- These constraints do not eliminate one of both symmetric forms for those nonsymmetric configurations containing as many chosen cells in the upper (left) halfsquare as in the lower (right) half, or in the NE triangle as in the SW triangle. Admittedly there shouldn't be too many of them ...

  It should be possible to construct more complicated constraints to discriminate even these cases, but we feel their contribution will be so marginal that they will probably be of no use.

2.3 Valid inequalities

We now move to the determination of valid inequalities. Several families of such can be obtained by considering hexomino subsets of the board as described below.

2.3.1 2×3 subrectangle

In [2] a family of valid inequalities is suggested, stating that in every 2×3 subrectangle of the square at least 2 cells are used. Inclusion of these inequalities — together with the two symmetry breaking constraints (4) and (5) — resulted in an IP model, solvable by CPLEX after enumeration of 5605 nodes. That these constraints are quite effective may already be seen from the LP relaxation, which has optimal value 21 when $n = 8$, a quite important improvement over the value 12.8 found without them.

The reason behind the validity of these constraints may be seen in two ways. Consider any 2×3 subrectangle H.

- Deleting (using) any single cell of H still leaves room for a (B shaped) pentomino, which is unfeasible. Therefore in any case a second cell of H should be used (at least).

- Consider all pentomino placements included in H. There are 6 of them, all B shaped (see table 1), each one being obtained by deleting one cell from H. To each corresponds an exclusion constraint (2), which may be written as the family

$$\forall r \in H : \sum_{i \in H \setminus \{r\}} x_i \geq 1$$

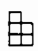
Which yields by summation over all $r \in H$

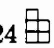
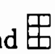
$$\sum_{r \in H} 5x_r \geq 6$$

and after division by 5 and rounding up the right hand side 6/5 to the nearest integer 2 one obtains the sought valid inequality

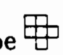
$$\sum_{r \in H} x_r \geq 2 \qquad (7)$$

This second derivation clearly shows that constraint (7) is stronger than the combined effect of the 6 exclusion constraints. In fact it is even strictly stronger than each of them separately : using $x_r \leq 1$ for each choice of $r \in H$, (7) implies immediately each of the 6 corresponding exclusion constraints.

Since any shape B or C pentomino is evidently included in a 2×3 subrectangle, it follows that when constraints (7) are added to the formulation there is no need anymore to also include those constraints (2) excluding any of the  or  pentomino placements.

In other words, the $12n^2 - 36n + 24$  and  placements excluding constraints are advantageously replaced by $2n^2 - 6n + 4$ (2×3)-subrectangle constraints. The resulting IP formulation will not only be combinatorially equivalent to the original formulation, but even have a strictly stronger continuous relaxation.

2.3.2 The &-hexomino

Consider now the hexomino of shape  which we will call the &-hexomino, a placement of which we denote by H.

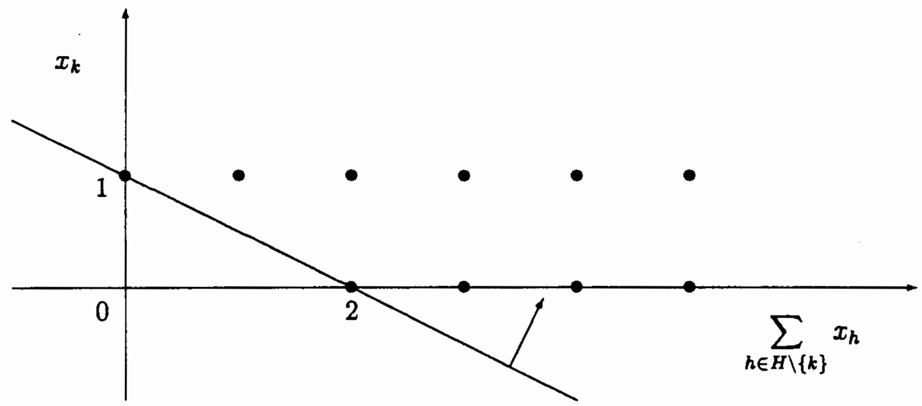
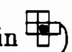
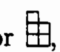


Figure 1: Deriving the &-hexomino constraint

- Deleting any of the 5 border cells (all except the ■ in ) from H yields a pentomino of either shape , , or , while H's kernel cell k (indicated as ■) disconnects H when deleted. Therefore in any feasible configuration, when this kernel cell is free, then at least two of the border cells must be used.

- In algebraic terms, for each border cell of H we have an exclusion constraint (2),
$$\forall h \in H \setminus \{k\} : \sum_{i \in H \setminus \{h\}} x_i \geq 1$$
which yields by summation over all $h \in H \setminus \{k\}$
$$\sum_{h \in H \setminus \{k\}} 4x_h + 5x_k \geq 5.$$
In case $x_k = 0$ this yields $\sum_{h \in H \setminus \{k\}} 4x_h \geq 5$ and after division by 4 and rounding up the right hand side 5/4 to the nearest integer 2 one obtains
$$\sum_{h \in H \setminus \{k\}} x_h \geq 2. \tag{8}$$

The inequality (8) may, however only be invoked under condition that $x_k = 0$, and should be inoperative when $x_k = 1$. This is obtained by way of the following constraint
$$\sum_{h \in H \setminus \{k\}} x_h + 2x_k \geq 2 \tag{9}$$

which is most easily derived from the graph in figure 1 in which the combinations of values which are to remain feasible are shown as •'s.

This yields a valid inequality for each of the $4(n-2)^2$ possible placements of a &-hexomino on the board. The derivation above, and the fact that a strict upper rounding was performed suggests that we have derived genuine cutting planes. This is strongly emphasized in Section 2.4 where it is shown that all the derived valid inequalities are in fact facet generating.

We may therefore expect that adding these constraints will improve the LP approximation bounds, and should decrease the size of the branch and bound tree. It turns out that the LP-relaxation optimal value for $n = 8$ remains equal to 21, so no further gain is observed at the highest tree-level over what was achieved by the (2×3)-rectangle constraints.

These &-hexomino constraints also do not seem to render some of the pentomino-exclusion constraints redundant.

2.3.3 4- and 3-border hexominoes

The valid inequality derived in previous section may now easily be generalised to other hexomino shapes.

A *border cell* of some polyomino is any of its cells which can be deleted without disconnecting. In other words, deleting a border cell from a polyomino still yields a polyomino. Non-border cells will be called *center cells.*

Let us call a *p-border hexomino* one which has exactly p border cells. Such a hexomino may also be defined as being the union of exactly p distinct placements of pentominoes.

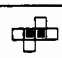
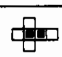
Observe that the (2×3) rectangle is the only 6-border hexomino, while the &-hexomino is the only 5-border one. Of the 35 hexominoes 8 are 4-border hexominoes (see table 2) and 12 are 3-border hexominoes (see table 3). In these tables the kernel cells are indicated by ■ (the dots are for later use). All of the 13 other remaining hexominoes are 2-border.

Consider any p-border hexomino placement H $(p = 4, 3)$, with set of border cells B and set of kernel cells $K = H \setminus B$. Note that $\#H = 6$, $\#B = p$, $\#I \setminus H = n^2 - 6$ and $\#K = 6 - p$.

Similar to previous section we may now say that when all cells of K are free then at least two of all border cells must be used. Indeed in case all of K remains free, and using only one border cell of H leaves a full pentomino free, which is not feasible. In other words, **if** $\sum_{k \in K} x_k = 0$ **then** $\sum_{b \in B} x_b \geq 2$, which, as in previous section, is expressed by the constraint

$$\sum_{b \in B} x_b + 2 \sum_{k \in K} x_k \geq 2 \tag{10}$$

Table 2: list of all 4-border hexominoes

Facet generating				
Shape	# forms	box	#plac/form	#placements
	8	(3,3)	$(n-2)^2$	$8n^2 - 32n + 32$
	8	(4,2)	$(n-3)(n-1)$	$8n^2 - 32n + 24$
	4	(3,3)	$(n-2)^2$	$4n^2 - 16n + 16$
	4	(4,2)	$(n-3)(n-1)$	$4n^2 - 16n + 12$
	4	(4,2)	$(n-3)(n-1)$	$4n^2 - 16n + 12$
	4	(4,3)	$(n-3)(n-2)$	$4n^2 - 20n + 24$
	4	(4,3)	$(n-3)(n-2)$	$4n^2 - 20n + 24$
Total				$36n^2 - 152n + 144$
Not facet generating				
	8	(3,3)	$(n-2)^2$	$8n^2 - 32n + 32$
Total				$8n^2 - 32n + 32$

Note that for $p = 6$ and $p = 5$ these constraints are exactly the (2×3)-rectangle and &-hexomino contraints, since $K = \emptyset$ and $K = \{k\}$ respectively.

Table 2 shows that we have $44n^2 - 184n + 176$ such 4-border hexomino constraints, and according to table 3 there are $84n^2 - 400n + 432$ 3-border constraints. For $n = 8$ this means 1520 4-border and 2608 3-border hexomino constraints.

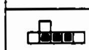
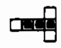
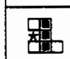
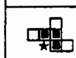
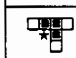
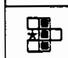
In the section 2.4 it is shown that the 4-border and 3-border hexominoes in the top part of these tables yield constraints which are facet generating.

2.3.4 2-border hexominoes

When one tries to apply the technique of previous section to some 2-border hexomino, the resulting constraint turns out to be exactly the sum of the two constraints of type (2) corresponding to the two pentominoes included in the 2-border hexomino.

Hence 2-border constraints are of no interest, since they would simply be redundant.

Table 3: list of all 3-border hexominoes

Facet generating				
Shape	# forms	box	#plac/form	#placements
	8	(4,3)	$(n-3)(n-2)$	$8n^2 - 40n + 48$
	8	(4,3)	$(n-3)(n-2)$	$8n^2 - 40n + 48$
	8	(4,3)	$(n-3)(n-2)$	$8n^2 - 40n + 48$
	8	(4,3)	$(n-3)(n-2)$	$8n^2 - 40n + 48$
	8	(5,2)	$(n-4)(n-1)$	$8n^2 - 40n + 32$
	4	(4,3)	$(n-3)(n-2)$	$4n^2 - 20n + 24$
	4	(5,2)	$(n-4)(n-1)$	$4n^2 - 20n + 16$
Total				$48n^2 - 240n + 264$
Not facet generating				
	8	(3,3)	$(n-2)^2$	$8n^2 - 32n + 32$
	8	(4,2)	$(n-3)(n-1)$	$8n^2 - 32n + 24$
	8	(4,3)	$(n-3)(n-2)$	$8n^2 - 40n + 48$
	8	(4,3)	$(n-3)(n-2)$	$8n^2 - 40n + 48$
	4	(3,3)	$(n-2)^2$	$4n^2 - 16n + 16$
Total				$36n^2 - 160n + 168$

2.4 Facets

Let us call the convex hull of all pure 0-1 feasible points of the pentomino-exclusion problem given by (1)-(3) the *pentomino-exclusion polytope.* It turns out that many of the defining and valid inequalities derived in previous section in fact define facets of this polytope.

Indeed the classical study of facets for the general covering polytope by Balas and Ng [1] gives all the results necessary to obtain the following conclusions.

Theorem 1
The following properties hold for the pentomino-exclusion polytope:

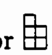
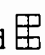
1. *It is fully dimensional, i.e. of dimension n^2.*

2. *All the defining pentomino-exclusion (or* covering*) constraints (2), except B and C, are facet generating.*

3. *The valid inequalities (7), (9) and (10), respectively generated by the (2 × 3)-rectangle , the &-hexomino, the seven 4-border hexominoes in the top-part of table 2 and the seven 3-border hexominoes in the top-part of Table 3 are the only facet generating p-border hexomino-constraints.*

Proof

The first statement follows from the fact that at least 2 (in fact exactly 5) variables x_i appear in every covering constraint (2) (see [1], 1., p58).

According to [1], 5., p59, a covering constraint (2) $\sum_{i \in P} x_i \geq 1$ $(P \in \mathcal{P})$ defines a facet if and only if two properties hold: first, that no other $P' \in \mathcal{P}$ exists with $P' \subset P$, which is evident here since all $P \in \mathcal{P}$ have the same cardinality, and secondly, that for each $k \notin P$ there exists some $k' \in P$ which lies in every $P' \in \mathcal{P}$ with $k \in P' \subset P \cup \{k\}$. In our context this last property means that the hexomino obtained by adding any additional cell to $\mathcal{P}$ can be disconnected by taking out just one other cell. For most pentominoes this particular disconnecting cell may be chosen independently of the added cell, as shown by a dot in table 1. For the W pentomino a disconnecting cell always exists, but not a fixed one. For  and  one may obtain a full 2 × 3-rectangle by adding just one cell, which cannot be disconnected by deletion of a single cell.

All p-border hexomino-constraints have coefficients in $\{0, 1, 2\}$, hence are of the particular type studied in [1], and can therefore be tested for being facet-defining thanks to their theorem 2.6. We restate this theorem below using our notation, i.e. H for some hexomino, $K \subset H$ its kernel and $B = H \setminus K$ its border, and discuss the implication of each part to our particular pentomino setting.

1. H is minimal, i.e.

   - $H = \cup\{ P \in \mathcal{P} \mid P \subset H \}$ and $K = K_H := \cap\{ P \in \mathcal{P} \mid P \subset H \}$, both of which evidently hold for all hexominos H.
   - if $H' \subset H$ and $K_{H'} = K_H$ then $H' = H$, which is evident also in our setting, since the only strict subsets of a hexomino H large enough to hold a pentomino are itself pentominoes P, for which $K_P = P$ which can never equal $K_H = K$.

2. A pair $\{j, j'\} \subset B$ is called a 2-cover of H if every $P \subset H$ meets $\{j, j'\}$. The second condition now states that the 2-cover graph on B must contain an odd cycle in each of its components.

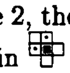
   In our setting every pair in B (such exist as soon as $p \geq 2$) is evidently a 2-cover, so the corresponding 2-cover graph is a complete graph on p nodes, which contains an odd cycle as soon as $p \geq 3$.

3. For some $k \notin H$ let $T(k) = \{ P \in \mathcal{P} \mid k \in P \subset H \cup \{k\} \}$. The last condition states that for every $k \notin H$ with nonvoid $T(k)$

   - either some $k' \in K$ belongs to all $P \in T(k)$
   - or some pair in B meets every $P \in T(k)$ and every $P \subset H$

   Observe in our setting that the only $k \notin H$ with nonvoid $T(k)$ are the neighbour cells of H, i.e. which have at least one edge in common with H, or, in still other words, addition of which to H yields a heptomino. Hence we may translate the condition as follows :

   - In every heptomino containing H all subpentominos can be excluded by way of either one kernel cell or two border cells of H.

   This last property is easily checked. E.g. for the (2×3)-rectangle, the &-hexomino and the first four 4-border hexominoes in the top-part of table 2, the fixed pairs of border cells indicated by a dot respectively in  , in  and in table 2 will disconnect any heptomino obtained by adding one neighbour cell. For the other hexominoes in the top-part of tables 2 and 3 the property can also be checked, although not with a fixed pair in B.

   For each hexomino in the bottom-part if these tables a $\star$ indicates a neighbour cell which violates the property, showing them not to be facet defining.

∎

A direct proof, using only elementary notions of linear algebra, of some of the above results can be found in [5].

2.5 Final observation

Following up on an earlier version of this work [5] Glerup and Larsen [3] have obtained many more results. They observed that every n-border hexomino constraint ($n > 2$) is combinatorially equivalent to the set of pentomino exclusion constraints for those pentomino placements that are covered by it. They study what other constraints (involving larger subsets of the board) also have this property, and how to exploit this in order to obtain the most economical formulations in terms of number of constraints. This enabled the solution of the pentomino exclusion problem on many rectangular and torus shaped boards, of dimensions for which the total number of constraints (even when restricted to the facet-generating ones) was prohibitive.

3 The pentomino spanning problem

3.1 Introduction

Let us define a pentomino-shape (often just called a pentomino) as the set of all its placements, i.e. as a set of subsets of the board.

Placing several pentominoes on the board implicitly means that the corresponding placements are pairwise disjoint subsets of the board.

A pentomino(-shape) is excluded if it cannot be added to the board without overlapping other already placed pentominoes. Therefore a set of placements of pentominoes (or configuration) spans the board when it cannot be extended to include more pentominoes.

3.2 A general set spanning problem and an IP formulation

Formalizing these observations we define the following more general problem.

Consider a finite set I (*the board*) and a set $\mathcal{P}$ of subsets of I (*the placements*). $\mathcal{P}$ is partitioned into a set $\mathcal{S}$ of *shapes* , where each $S \in \mathcal{S}$ consists of the different possible placements of shape S.

A *configuration* is a set of placements, at most one for each shape, which are pairwise disjoint.

A configuration *spans* the board if it is maximal for inclusion, i.e. there is no configuration that contains additional placements.

The *general set spanning problem* is to find a spanning configuration of minimum size (cardinality).

We introduce a binary variable $y_P \in \{0,1\}$ for each placement $P \in \mathcal{P}$, which indicates whether P is an element ($y_P = 1$) or not ($y_P = 0$) of the configuration.

Clearly the objective is to minimize the number of placements in the configuration, i.e.

$$\min \sum_{P \in \mathcal{P}} y_P \tag{11}$$

To be feasible the configuration must satisfy two conditions:

1. **For each shape the configuration contains at most one of its placements**

$$\sum_{P \in S} y_P \leq 1 \quad \text{for each } S \in \mathcal{S} \tag{12}$$

   in other words: the variables y_P ($P \in S$) form a special ordered set (of type 1) for each $S \in \mathcal{S}$.

2. **The placements are pairwise disjoint**
   which may be restated as: every point $i \in I$ of the board may be covered at most once by the placements. This is expressed by the inequalities

$$\sum_{P \in C_i} y_P \leq 1 \quad \text{for each } i \in I \tag{13}$$

   where $C_i = \{P \in \mathcal{P} \mid i \in P\}$.

Finally we must also express that the configuration should **span the board**. In other words we must express as constraints that 'if a shape S is not represented in the configuration, then none of its placements can be added to it'.

First note that for any shape S the expression $\sum_{P \in S} y_P$ can take on only values 1 or 0, meaning that S is represented in the configuration (has been *placed*), or not, respectively. We therefore have implicit binary variables (see [4])

$$z_S = \sum_{P \in S} y_P$$

stating whether S is placed or not.

Secondly, observe similarly for each $i \in I$ we have an implicit binary variable

$$x_i = \sum_{P \in C_i} y_P$$

which states whether the point i is covered by the configuration, or not.

What must be expressed now is the following property :

for any $S \in \mathcal{S}$ and any $P \in S$:
if S has not been placed,
then not all $i \in P$ may be uncovered

or,

if all $i \in P$ are uncovered ($\forall i \in P : x_i = 0$),
then S must have been placed (elsewhere) ($z_S = 1$).

This is expressed by the logical implication constraint (see [4])

$$\sum_{i \in P} x_i \geq 1 - z_S$$

The correctness of this constraint can be checked also as follows : in case $z_S = 0$ we obtain the constraints used in [2] to prohibit the placement P, while in case $z_S = 1$ the inequality becomes redundant.

In other words, and in terms of the original variables y_P, the spanning property is expressed by the constraint set

$$\sum_{i \in P} \sum_{Q \in C_i} y_Q + \sum_{Q \in S} y_Q \geq 1 \quad \text{for each } S \in \mathcal{S} \text{ and } P \in S \tag{14}$$

This completes the IP formulation, consisting of (11), (12), (13), (14).

3.3 Observation

It must be observed that, similarly to minimum covering problems, usually there will be many optimal solutions. This is a consequence of the fact that at the one hand the objective only counts the number of used placements, and so can only take on values between 0 and $\#\mathcal{S}$, while at the other hand the number of different configurations is of much higher order in $s = \#\mathcal{S}$ — when all S have similar cardinality p, and only taking the first constraints into account we may have $O(p^s)$ configurations. It follows that the same objective value will usually appear for large numbers of configurations.

For this reason B&B methods will probably work badly on this standard formulation. It is therefore necessary to devise and include additional constraints aimed at reducing this degeneracy. It is not clear how this might be achieved on the general problem described above, but for more specific instances, like the pentomino on a square board case several things may be done as explained in the section 3.5.

3.4 The pentomino spanning problem in particular

According to Table 1 the formulation developed above contains $63n^2 - 240n + 196$ binary variables grouped into 12 sets (shapes) out of which at most one may be chosen. Therefore an upper bound on the total number of feasible solutions is n^{24} (in fact exactly the product of all values (+1) in the last column of the Table): just disregard any of the non-SOS constraints.

The total number of constraints are:

# Forms / shape	type (12)	12
Disjointness	type (13)	n^2
Spanning	type (14)	$63n^2 - 240n + 196$

In other words for the standard 8×8 board we have 2308 binary variables and 2384 constraints.

3.5 Symmetry breaking constraints

An important way of limiting the number of equivalent solutions is to take symmetry into account. Evidently any solution yields equivalent solutions by any symmetry of the square. What should be done is to accept only one representative among these symmetric solutions. A way of obtaining this reduction is as follows.

Consider a shape S which is placed in the configuration, and select one particular symmetric form of it. By applying symmetries to the square board it is always possible to bring this particular shape S in the chosen position. Therefore accept only this particular form to appear, thus forbidding all other equivalent symmetric solutions.

Of course this reduces most when S has many forms. Therefore we preferably apply this method to the shapes with 8 forms, if available.

So let us number the shapes in the table from S_1 to S_{12}, sorted on number of forms (and as a second key the number of placements) — Table 1 was already sorted this way. We want to make use of the first shape, i.e. the S_i $(1 \leq i \leq 12)$ which is placed ($z_{S_i} = 1$), but none of the previous ones has been placed ($z_{S_j} = 0$ for all $j = 1, \ldots, i-1$). And for this shape S_i we accept only one of its forms (e.g. the one shown in the table), i.e. we impose that all other forms of this shape may not be placed. Denoting by S_i^* the set of all placements of the chosen form we have therefore the following constraints

'if $z_{S_i} = 1$ and $z_{S_j} = 0$ for all $j = 1, \ldots, i-1$,
then $y_P = 0$ for all $P \in S_i \setminus S_i^*$'

or (see [4])

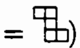
$$y_P \leq 1 - z_{S_i} + \sum_{j=1}^{i-1} z_{S_j} \quad \text{for all } i = 1, \ldots, 12 \text{ and } P \in S_i \setminus S_i^*$$

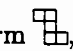
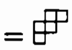
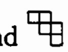
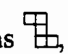
or, after plugging in the expressions of the $z_S = \sum_{Q \in S} y_Q$:

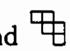
$$y_P + \sum_{Q \in S_i} y_Q - \sum_{j=1}^{i-1} \sum_{Q \in S_j} y_Q \leq 1 \quad \text{for all } i = 1, \ldots, 12 \text{ and } P \in S_i \setminus S_i^* \quad (15)$$

The total number of symmetry breaking constraints is $51n^2 - 194n + 159$, which is the table's last column sum minus the one-to-last column sum. For $n = 8$ this leads to an extra 1935 constraints.

This number may be reduced to only 12, by replacing the full set for each shape by one constraint obtained after summing all left hand sides of all constraints of type (15) for every $S \in \mathcal{S}$. However, one may expect these relaxed symmetry breaking constraints to be much less performant ...

Note that this symmetry breaking is slightly incomplete. Indeed, when none of the shapes with 8 forms ($i = 1, \ldots, 5$) have been placed in the configuration, only the symmetries of the first shape placed, ($S_6 = Z =$ ) will be broken, and this does not represent the full symmetry group of the square.

Indeed, choosing for S_6^* all the placements of the form , this still allows any feasible solution to be rotated over an angle of π, which may yield an equivalent feasible solution. Excluding this calls for a modification of the symmetry breaking constraints for $i = 6$, expressing that additionally among the 4 forms of the next shape ($S_7 = W =$ ) only 2 are allowed excluding their π turned forms, e.g. allow only the two forms  and . This means we need constraints that impose that as soon as none of the shapes with full symmetry group $(1, \ldots, 5)$ are used, then only the placements of forms ,  and  are allowed for shapes Z and W.

Defining $\bar{S}_7$ as the set of all placements of forms  and , we may write the additional constraints

$$y_P \leq (1 - z_{S_6}) + (1 - z_{S_7}) + \sum_{j=1}^{5} z_{S_j} \quad \text{for all } P \in S_7 \setminus \bar{S}_7$$

Many more similar symmetry breaking constraints may be written, e.g. when shape Z is the first to be used, but W is not used, or when W is the first shape to be used, etc. These are left to the reader.

It must be clear, however, that these additional constraints will often not be very powerful, since they only have any chance to be activated under very stringent conditions: when many of the shapes are not used. And this will happen only on smaller boards, where the problem tends to be easier, so probably the corresponding problem instance will be already easily solvable without them.

3.5.1 Valid inequalities

For each valid inequality of p-border type ($p = 6, 5, 4, 3$) for the pentomino-exclusion problem, we may deduce a corresponding valid inequality for the pentomino spanning problem, as follows.

Consider some p-border hexomino placement H with border B and kernel $K = H \setminus B$. Let us denote by $\mathcal{S}(H)$ the set of pentomino shapes of which a placement is contained in H, i.e.

$$\mathcal{S}(H) = \{S \in \mathcal{S} \mid \exists b \in B : H \setminus \{b\} \in S\}$$

We can then state the following property which should hold for any feasible configuration :

> If none of the (placements of the) shapes in $\mathcal{S}(H)$ is used,
> then the p-border constraint (10) corresponding to H must hold.

Indeed, if this is not true, then H would contain a "hole" large enough to fit in a pentomino of some shape in $\mathcal{S}(H)$ that is still free, which is forbidden for spanning configurations.

In other words we must have the implication :

> If $\sum_{S \in \mathcal{S}(H)} z_S = 0$,
> then we must have $\sum_{b \in B} x_b + 2 \sum_{h \in K} x_h \geq 2$.

and this may be expressed in a similar way as shown in Figure 1 by the constraint

$$\sum_{b \in B} x_b + 2 \sum_{h \in K} x_h + 2 \sum_{S \in \mathcal{S}(H)} z_S \geq 2 \tag{16}$$

or, using the definition of the implicit variables x_i and z_S,

$$\sum_{b \in B} \sum_{P \in C_b} y_P + 2 \sum_{h \in K} \sum_{P \in C_h} y_P + 2 \sum_{S \in \mathcal{S}(H)} \sum_{P \in S} y_P \geq 2 \tag{17}$$

This leads to a quite large number of supplementary constraints.

4 Final observations

4.1 Exclusion vs spanning

Note that the pentomino exclusion problem in [2] is of a quite different nature than the spanning problem.

In the exclusion problem the set of shapes which may be used to block (the monominoes) and the set of shapes to be blocked (the pentominoes) are fixed in advance, and even disjoint (but this latter can easily be lifted in a generalized statement). In the spanning problem these sets are one and the same and it is the solution that has to make the choice which shapes are blocking and which are blocked, implying many more decisions to take.

This explains why the spanning problem looks (and most probably is) so much harder than the exclusion problem.

4.2 Generalisations

4.2.1 Polyominoes

Evidently, the pentomino exclusion and spanning problems discussed here are directly generalisable, together with both the symmetry breaking constraints and the valid inequalities to higher order polyominoes and other shapes of boards.

4.2.2 Generalised exclusion

Also general versions of the exclusion problem can easily be defined, in which the board is an abstract set, and the excluded placements are just general subsets.

One may also generalise the individual cells used for exclusion to a fixed set of subsets, as being the blocks which are allowed to be used in order to exclude the forbidden ones. This was done in [3] for pentomino exclusion by dominoes.

4.3 Further generalisations of spanning

One might wish to consider spanning problems allowing for more than one placement of some shape.

The temptation to limit the number of such placements by changing the right hand side of the constraints (12) must be resisted. Indeed it is only when this right hand side equals 1 that the expressions $\sum_{P \in S} y_P$ are implicit binary variables. But when these are allowed to take on other values than 0 or 1, then our spanning constraints (14) are incorrect.

There is in fact a quite simple way to allow for more than one placement of some shape. Observe that the statement of the general spanning problem in section 3.2 allows for different shapes with common placements. Therefore it also allows for several copies of a same shape, which is exactly what was sought here.

Observe, however, that this introduces even more degeneracy in the problem, which may be handled by some new symmetry breaking constraints, now in order to break the permutation symmetry between the copies of one and the same shape. This type of symmetry breaking was already discussed in [4].

Acknowledgement We thank J.Fliege for his patient listening to and reading of several successive versions of this material.

References

[1] E.Balas and S. M. Ng, On the set covering polytope: I. All the facets with coefficients in $\{0, 1, 2\}$, *Mathematical Programming*, 43 (1989), 57-69

[2] R.A.Bosch, A pentomino exclusion problem, *Optima*, 60 (dec 98), 3

[3] N. Glerup and P.B.Larsen, The pentomino exclusion problem, Working paper, Department of Computer Science, University of Copenhagen, december 1999, 103p. (email: nanna@diku.dk, pebl@diku.dk).

[4] F.Plastria, Formulating logical implications involving 0-1 variables, Working paper BEIF/103, Department of Management Informatics, Vrije Universiteit Brussel, 1998.

[5] F.Plastria, On IP formulations for the pentomino exclusion and spanning problems, Working paper BEIF/110, Department of Management Informatics, Vrije Universiteit Brussel, May 1999.

[6] Wolsey, L.A., (1998), *Integer programming*, Wiley-Interscience.