OPTIMIZED PREFERENTIAL BIDDING SYSTEMS
Models and Implementations

by Daniel Tumpson
Crewing Solutions LLC
October 27, 2005

Introduction:

In a conventional PBS, each crewmember sends in his bid (i.e. preferences) without knowing the preferences of other crewmembers. He is bidding "blind". As a consequence, conflicts result: some trips are desired by (i.e. fulfill the preferences of) too many crewmembers and some are desired by no crewmembers. In the post-bid optimization process, the global optimizer attempts to reconcile conflicting bids (as represented by the final set of conflicting preferences). The optimum solution, even if it could be reliably extracted from the huge set of possibilities (using, say, LP and column generation methods), is constrained in quality by the conflicting preferences.

SmartPref, in contrast, transforms the process of bid submission into a "pre-optimization" process, wherein crewmembers are made aware each time they bid of the status of other crewmember bids: e.g. while junior crewmembers are allowed to bid generic preferences that potentially apply to all trips, they are not allowed to bid for specific trips assigned to more senior crewmembers. This feedback from other crewmembers combines with a bidding crewmember's preference information to order the trip choice list and to generate the best quality crew roster available at the time he bids. If his preferences produce a poor roster (e.g. because of conflicts with senior crew), the bidding crewmember may change his preferences until the roster he likes the best and which is also compatible with other crew bids is generated. The resulting bid, then, is "locally optimal" in that both the bid roster and the bid preference parameters are the best they can be given the present state of the solution and the seniority of the crewmember. To insure that this feedback is available from the beginning of the bid period, a reliable initial solution is generated from the standby bids of each crewmember.

By the end of the bid period, the bidding crew members have not only generated what they consider to be the best rosters that are compatible with other crewmembers, but also the preferences that they consider to be the most realistic given feedback from other crewmembers. Since the interactive bid process generates both a solution (set of rosters) and a crew-approved set of preferences, the interactive bid process simultaneously optimizes both crew preferences and crew rosters. The post-bid build/optimization process can improve the solution to some extent, e.g. by working open time trips into the solution and through global trip swap optimization, but the initial build and preference parameters have been adjusted by the crew themselves to fit together with minimal conflict.

No matter what post-bid build/optimize process is in place, an improved solution will be obtained over a non-interactive process due to the crew-approved parameter optimization that occurs during the interactive bidding process.

Solving the PBS Problem: General Considerations:

Crewing Solution's SmartPref Preferential Bidding System is designed to produce solutions for monthly crew assignments ('lines') based on crewmember preferences, or is what has become known as a *Preferential Bidding System* (PBS).

It aims to replace the current *bid line* systems that are mostly in use by North American carriers today. In the bid line approach, trips are assigned to *generic* lines, which do not take into account individual crew preferences or constraints, such as carry-in, vacations, and other non-trip work activities (e.g. training). The generic set of lines are then put up for bid to the crew and awarded in seniority order. The bid line solutions have the obvious defect that, since they were created without reference to specific crew constraints, they do not integrate well with the crew specific non-trip work and rest activities. This lack of integration is often exploited by crew members in that they may seek to be awarded lines that conflict with their predetermined activities to maximize pay and credit. The conflicting trips must be assigned to other crewmembers resulting in a wide range of utilization (under and above a desired and expected utilization range). Clearly this uneven allocation is costly to the airlines and in today's world of dwindling profits, airline management is eager to streamline operations and utilize their resources as best they can. PBS has therefore become a natural alternative to bid lines and is being endorsed more and more by both crewmembers and management in the effort to improve financial and operational performance. Also it should be noted that the idiosyncrasies of bid lines and its awkward handling of schedule continuity is long overdue for a complete overhaul. PBS is the logical next step.

Management will be quick to point out that the PBS method that is adopted matters little to them as long as the upper and lower bound of utilization are met. Most PBS systems on the market meet these requirements, so that the choice of PBS is, for the most part, crew driven It is therefore crucial for the crews to understand how the PBS they endorse works and why it is demonstrably better than others. The remainder of this paper will contrast the various approaches available today and explain the novel (patented) approach offered by Crewing Solutions' SmartPref.

In any type of PBS, crew preference information is conveyed by each crewmember to the scheduling body via a *bid* which must be submitted within a *bid period*. This preference information is then used to create the PBS line solution. However, there are different approaches to structuring the PBS bid entry which will result in different levels of solution quality and in different time scales for accomplishing the solution.

For example, the standard approach (used by Carmen Systems and other vendors) is to collect bids *non-interactively* from crewmembers during the bid period and then to use this bid information to generate a global solution using a combination of linear optimization (LP) methods and heuristics (which may be combined, e.g., into a composite method such as *column generation*). Even though the bid may be submitted via internet or other type of on-line network, it is non-interactive in the sense that no projected realization of the bid is returned. While some systems (such as Atlas Bid at Northwest) may offer planned-inventory availability and search functions, they do not project any results based on the overall demands of participating crewmembers.

Although the PBS solution can be linearly modeled and so is, in theory, subject to LP methods, the problem size (involving the number of possible lines from which to choose the solution set) is so vast that LP methods alone cannot be used to obtain a near-optimal solution. Heuristics (such as those used in column generation) have been developed for using LP solution information to effectively reduce the solution space (i.e. the line choice set) to a manageable size, but they may suffer from several weaknesses: (1) the heuristics cannot be proven to evolve the solution toward global optimality; (2) the total time needed to derive the solution can be very large (even at today's computer speed), which can be a problem in recovery and other situations where quick solutions are needed, and (3) clashes in crew preferences may lead to solutions which are unsatisfactory compromises, generally due to heuristic assumptions unseen by the bidders.

In contrast, the Crewing Solutions SmartPref PBS accomplishes *preference parameter optimization* through *interactive bidding* which takes advantage of crew seniority to create a solution which is much more likely to be free from crew preference conflicts and which can be generated and updated in a matter of minutes. This solution speed is vital to the SmartPref strategy, which allows a crew member to bid interactively with the PBS via an internet accessible graphical user interface (GUI) thereby registering his or her preferences *in relation to the preferences of crew who have previously bid*. Unlike the standard approach, where the crew are bidding blind, the SmartPref interactive bidding allows crew members to know at the time of their bid what trips are realistically available and/or what days off they are likely to hold. In addition, crew may interactively update their bids as often as they want during the bid period, so an ever more accurate adjustment of preferences to availability may be accomplished.

The SmartPref interactive bid phase, then, may be seen to be an optimization of all the choices among the available trips and days off among the participating crew members and, by the end of the bid period, these preferences, as represented by preference parameters for all bidding crew will have been effectively optimized. These parameters, and the resulting short list of trips assignable to each crew member, can then be used to generate a final solution which is of high quality and relatively free of bid conflicts.

Most importantly, the repeated adjustment of preferences by bidding crew will also result in ever more realistic expectations, so that the final assignments will be much more likely to satisfy their expectations. This will lead to a genuine cooperation among crew members and between crew and management resulting in line assignments which satisfy both the crew's preferences and the airline's productivity requirements.

The PBS Line Solution problem can be reduced to an *integer programming* (IP) problem which can be approached with linear programming (LP) methods (see Appendix 1), including column generation, which appears to be the method of choice for a number of major crew scheduling solution providers, including Carmen Systems and other vendors. The column generation approach to solving the lines problem is discussed in Appendix 2, emphasizing the sensitivity of the method to the choice of heuristics which determine the order in which columns are generated and added to the evolving choice set. (The more technical aspects of the various approaches have been included in appendices so as to make the reading and understanding of the methods discussed accessible to all.)

And finally, we will describe the patented Crewing Solutions SmartPref PBS system, discuss how it addresses the problems of choice set size, and explore the advantages of SmartPref's unique interactive bidding process. As we shall see, by first working with crew members interactively to define preference parameters (through their bid) to insure in advance that a large number of acceptable lines can in fact be built, and by taking advantage of crew seniority ordering in the build and optimization process, the SmartPref method actually increases the probability of obtaining a high quality solution very quickly.

Solving the PBS Problem: the Formal Approach:

It is assumed here that prior to the PBS process, the set of all flight legs to be covered in the scheduling period have been assigned to trips (pairings) and that the minimal scheduling object is a trip (a series of flight starting and ending at base) with a definite associated cost (credit).

The goal of PBS is to assemble a set of crew rosters (lines) that satisfy all roster constraints (e.g. no conflicts, credit within min/max bounds, minimal unassigned trips) and at the same time maximize crew satisfaction, and respect crew seniority.

In order to model the problem in a way that reveals its linear structure (and its susceptibility to LP and IP solution methods), we need to first define the choice set of scheduling objects from which the optimum solution may be extracted, and define the set of constraints and costs associated with these objects which can be linearly modeled. The obvious scheduling object for this purpose is the individual line assigned to a crewmember, i.e. the assignment of dated activities to the scheduling period for a single crew member. The goal will be to assemble a choice set (CS) of these lines and extract a solution from it using linear methods. To this end we must be able to assign a unique cost to each such line, and express line constraints such that they can be modeled linearly.

We first partition the possible constraints into two classes: "individual" or "horizontal" constraints, which apply to each line individually, and "collective" or "vertical" constraints, which apply to the entire solution set of lines. Examples of horizontal constraints included minimum and maximum block per scheduling period, minimum rest between trips, fixed activities, such as vacations and training periods, etc. For our present purposes, we will restrict our attention to the two principal classes of vertical constraints that must be satisfied by any optimum lines solution: (1) Each trip is assigned to one and only one line (2) the solution will contain one and only one line for each eligible line holder. These two are the only collective constraints that we will need to consider here to illustrate the approach. (See Appendix 1.)

Column generation (see Appendix 2) arose in response to the problem of applying LP methods to huge numbers of variables. In the case of both crew pairing and lines optimization, the choice sets of scheduling objects (trips and lines respectively) are huge: potentially on the order of billions. (See Appendix 4: estimate of PBS problem size.) In both cases, the scheduling object is the smallest to which legality can be ascribed and a cost can be associated, which are minimal conditions for an optimal solution to be extracted from a choice set with LP methods. The choice sets in both cases are way too large to be subjected to standard LP methods such as simplex. If one attempts to reduce the size of the choice set in some intelligent manner to some small and manageable subset, one runs the risk of throwing the baby out with the bath water. Unless that subset contains an optimal solution to the larger choice set, its solution will obviously be suboptimal.

As described in Appendix 2, column generation provides a method for including or excluding new lines into an evolving choice set - lines which will improve the optimal solution and that may be extracted from each successive choice set. However, there is no straightforward way to decide how best to add new candidates to the choice set. Obviously if the wrong ones are added, there will be little or no improvement, and this could occur over and over until the choice set is too big to solve with LP methods.

In the case of crew pairing optimization, very good results have been obtained, but the lines problem differs substantially from the pairing problem in that there are many more combinations of lines which are of comparable quality and therefore subject to being selected. In the case of pairings a minor variation (a different flight segment) will generally result in a change of cost, especially if some type of time away from base ratio is used, whereas in the lines problem there can be literally millions of lines that will satisfy a given criteria (e.g. all the lines that have between 72 and 80 hrs of credit and all Wednesdays off).

This suggests that solutions using heuristic methods which combine the simpler component scheduling objects, i.e. trips, may arrive more quickly at very good if not near optimal results. If such heuristics can indeed quickly generate very good solutions, this makes them a viable alternative to linear programming methods. Such trip-based heuristics are to be found in Crewing Solutions' patented SmartPref software.

Solving the PBS Problem Using Interactive Bidding and Component-Based Heuristics:

As alluded to above, there are a number of problems associated with the use of LP and column generation methods to solve the lines problem, including the huge problem size and the consequently long solution time. This has been recognized and addressed by Sellmann et al., 2002, who combine column generation with a heuristic tree search. Two observations emanate from this work. First, an evolving tree search of assigned trips is the solution build mechanism, i.e. the scheduling object involved is the trip, a line component. Although the method alone cannot produce global implications of optimality, it builds solutions from a smaller set of scheduling objects (trips) and is innately integer programming (IP) in nature. As a result, it can proceed more directly to feasible line solutions. Second, it relies on local search heuristics to resolve infeasibilities resulting from column generation and to improve the solution quality. In fact, the column generation component of this method proceeds from initial feasible solutions generated by heuristic tree searches.

As we shall see, when strict seniority is to be respected, a form of heuristic tree search can very quickly arrive at very good feasible solutions. This is at the heart of the SmartPref method, which creates line solutions in three phases.

In Phase 1, the crew members interact with the SmartPref system through an internet accessible graphical user interface (GUI). There, each crew member registers his or her preferences through multiple-choice questionnaires and are shown lists of available trips and corresponding lines which match their fixed activity structures (carry-in, holidays, training, etc.) and their stated preferences. (Since some trips match the preferences of crew members with higher seniority, they may be excluded from the list of trips available to crew members with lower seniority.)

If their preferences are too restrictive to permit a reasonable probability of being satisfied in the final line solution, they are encouraged to modify their preferences to increase this satisfaction probability. Through such interaction, the crew member is able to arrive at a set of potential lines that are both acceptable and likely to contain a line which is achievable in the final solution. It is assumed that each line in the choice set satisfies all horizontal line constraints and can be assigned a unique cost. As such, all lines that violate horizontal constraints are left out of the set, and a solution cost can be quantified. Obviously, this cost includes "hard costs" such as the sum of the credits of the trips therein, but also can be designed to included the "soft costs" associated with deviations from crew preferences, by establishing penalties for the kind of, and extent of deviation from, each preference. Crewing Solutions' SmartPref approach is to gather preference information in an initial interactive phase, in which crew members express their preferences and examine possible resulting options via the internet. This interactive phase allows multiple bid updates and building iterations that allows the bidder to adjust his demands based on the overall demands of crewmembers senior to him as explained below.

The bidding process is iterative and incremental. The crewmember bids by first searching for trips that he would like to fly. This search is made by clicking on various preference items (i.e., pairing number, day of week, departure time, etc.). The system returns to the crewmember the list of available trips matching the preference criteria. If the crewmember does not add these trips to his bid, these trips are automatically returned to the pool of available trips. If the crewmember likes the trip selection, he will add them to his bid at which time the system will build a line of time including this latest bid information. In one such step or a series of steps, the crewmember builds his bid and at every step, the system returns to him a line of time based on his bid. It is important to note that preferences are associated with an importance level given by the bidder to indicate how the penalties mentioned above will be treated. A bid generally has two components: (1) days off desired and (2) trips desired; these trips can be defined specifically or by general characteristics such as 3 day trips that layover at certain preferred cities. Obviously such trips may infringe on days off wanted and hence a level of importance, or priority, must be defined so that the penalties are assessed correctly. SmartPref not only externalizes this process, i.e. makes it visible to the bidder, but in addition produces lines that would result from the given choices.

Phase 2 is taking the information gathered in Phase 1 and in effect 'finalizing' it. It assumes that the bids as they stand are as they would be at bid closing. A global solution is derived insuring line completeness (i.e. credit within min/max bounds) and minimal open time (trips left unassigned). Lines produced in phase 2 use exactly the same logic as those tentatively produced in phase 1 with the only difference that certain global considerations are made (phase 3) to possibly improve the quality of the solution and in particular looking at improvements that could be made to junior lines by considering equivalent solutions for senior crew members that would benefit junior demands. Certain open time adjustments may be required depending on user defined limits on allowed open time. The Phase 3 process can be too time consuming to be contemplated in Phase 1. The results obtained at the conclusion of Phase 3 become the next basis for Phase 1. The process is repeated at fixed intervals or as often as feasible based on available processing (CPU) time.

Every eligible crewmember has a bid which translates into a bid matrix, an ordering of all the available trips in the sequence that will be submitted to a line builder from which an 'ordinal' line will be extracted. As the name indicates, the ordinal line is the result of adding to a line the trips in the order they were rated by the bidder. The ordinal line may be successful (within the stipulated limits of utilization) or not. If successful, it is retained as the first line in the possible line choices for the bid; otherwise the line builder generates 'combinatorial' lines. Combinatorial lines are derived by recombining or rearranging trip selection so as to insure completeness and maximum satisfaction. The recombination uses greedy substitutions algorithms that are very fast and exhaustive within the bounds set by the ordinal line and its successive improvements and a user defined parameter. Often an ordinal line cannot be improved and the best combinatorial line is then the ordinal line. Throughout the building process the awarding of trips is controlled by a 'stack control' that insures crew availability for the remaining unassigned trips.

Although the above is a simplification of the SmartPref heuristics, it gives the flavor of how a line is arrived at using trip-oriented manipulations, so that the complexity and number of possible lines is kept manageably small, and yet the twin objectives of achieving feasibility and optimizing cost and preference is attained.

Note that both the heuristics for bringing unassigned trips into the line solution and the binary-swap-optimization process are problem oriented: they are effective procedures for limiting the solution space search to neighborhoods of problem lines. Note also that the natural seniority hierarchy is exploited to speed up the build, the incorporation of unassigned trips, and swap optimization. The efficacy of the SmartPref line builder/optimizer strategy is evidenced by the fact that optimization that takes hours for CG methods to carry out (see e.g. discussions of the Carmen Crew Rostering System in Kohl and Karisch, 2004) can be accomplished in minutes with SmartPref.

The Advantages of Interactive Bidding:

Non-interactive PBS systems require that all crew submit their preferences without reference to the preferences of other crew members. These static preferences are then fed into the PBS builder/optimizer in the form of seniority and preference parameters which are then used to generate a global "assigned lines" solution. While non-interactive systems do utilize crew preference data to generate a global solution, they cannot avoid solution defects that arise from inevitable "clashes in preferences" which are characterized by sets of preferences for different crew members that would encourage the assignment of the same trip to two or more crew. Such preference clashes are symptomatic of non-interactive bid systems in that preferences are arrived at by crew members with no reference to the preferences of other crew. (See Appendix 3: Incompatibility of Non-Interactive Bidding and Parameter Optimization.)

In contrast, an interactive PBS has the potential to accomplish parameter optimization as part of the interactive process itself: when a bidder recognizes that his original preferences lead to an unacceptable line of time, he adjusts his preferences until the line is to his liking. The bidder's ability to adjust his preferences is thus the first advantage of the interactive PBS. The second, and complementary, advantage of interactive bidding is that it utilizes up-to-the-minute bidding information from other bidders to allow the crewmember to know what his possibilities are most likely to be under various bid scenarios. (In contrast, a non-interactive bidder has no information as to what the likely outcome of his bid will be: he is bidding blind.)

In order to accomplish an interactive bidding iteration, a crew member accesses his bid via the SmartPref GUI and inputs his preferences via menus which display and allow the bidder to choose in priority order the line characteristics (such as trip departure time, day of week, or op days) that are preferred or to be avoided. The SmartPref program then assembles an ordered list of trips which matches (or avoids) these characteristics and which are available, i.e. which have not been chosen by a senior crew member. In order for an interactive crew iteration to be accomplished, there must be an existing solution, from which the preferences of other crew can be extracted. (Prior to the first bid, the solution can either be initialized to a default set of standby bids, or to a null solution with no bids, trip lists, or line solution specifies and all trips initially in open time.)

Once the list of trips and resulting line choices have been made and approved by the bidder, the solution must be regenerated, taking into account changes to the solution imposed by the new bid. These changes are twofold: first, the bidder's new line choice must become part of the solution, and second, the impact of these new choices on the lines of other crew must be incorporated. Since the trips chosen by senior crew were removed from the bidder's choice list, there can be no effect of the bidder's new line choice on any line senior to him. But since the bidder can choose from any of the trips already chosen by junior crew, the solution must be updated to take account of the new bid on junior crew: any trips which are part of the bid must be removed from any junior lines. Any junior line from which such a trip is removed must then be rebuilt by the SmartPref program, i.e. a new set of line choices must be extracted from that junior crew member's list (minus the removed trip). In principle, this rebuilding of junior crew could be done recursively until all affected crew bids are successfully adjusted, so that the output of each completed crew bid is a complete and consistent solution, with trips possibly added to or removed from open time. But in order to avoid the complexities involving several crew bidding simultaneously and the possibly lengthy recursive build process, the rebuilding is instead performed at frequent periodic intervals in batch mode using a global build process.

Each interactive bid, then, can be considered an iteration within a larger solution build / parameter optimization process (with solution consistency imposed at frequent periodic time intervals via global batch rebuilds). At the end of the bid period, the parameter set is considered optimized and a tentative line solution has been built. As noted above, this process may leave trips in open time, which is then corrected in Phase II. Since this two-phase build process is strictly "top down" (in that no trip choices from more senior bidders and all trip choices from more junior bidders are included in a bidders choice list), no analysis of potential transfers of trips from senior to junior lines is carried out. Therefore, a final swap optimization process which investigates senior-junior trip swaps is carried out in Phase III.

Interactive bidding, then, works best if all crew members bid early in the bid cycle and continue to update their bids thereafter. This is due to the fact that the tentative assignment of trips to lines which occurs during interactive bidding relies not only on the preferences of the bidding crew member, but also on the *availability* of trips satisfying those preferences. This availability depends on the seniority and the bids of other crew members. If a senior crew member has chosen a trip, then it is not available to bidding junior crew members. In order for a junior crew member to have accurate information as to trip availability on which to base his bid, there must be an accurate record of senior bids at the time the junior crew member bids. This can only happen if at least one bid is made by every crew by a certain date, after which the complete and accurate record is deemed to be initialized. This is the primary reason why the use of standby bids to create an initial solution is highly recommended.

Thus the SmartPref interactive PBS provides the advantages of crew-driven parameter optimization in creating a solution which is more widely acceptable to all crew. When the bid closes, the solution builder/optimizer is able to take advantage of the resulting crew-optimized preference parameters, leading to a better final assigned lines solution.

Summary:

The Crewing Solutions SmartPref Preferential Bidding System was described and compared with line builders using column generation methods. Column generation is found to be a very powerful method for solving very large combinatorial optimization problems, such as the airline crew pairing problem, but suffers from the size of the problem dictated by the minimum size of the scheduling objects required for the use of column generation and other linear optimization methods. The use of interactive bid adjustment, seniority ordering, and build heuristics which depend on much simpler and less numerous scheduling objects, SmartPref permits the creation of high quality line solutions in a tiny fraction of the time required by linear methods.

Appendix 1: Solving the PBS Problem using Linear Programing techniques:

For our present purposes, we will restrict our attention to the two principal classes of vertical constraints that must be satisfied by any optimum lines solution ("OLS"): (a.) Each trip is assigned to one and only one line. (b.) The OLS will contain one and only one line for each line holder. These two are the only collective OLS constraints that we will need to consider here.

It is assumed that each line in the choice set (CS) satisfies all horizontal line constraints and can be assigned a unique cost. This simply means that all rosters (lines) that violate horizontal constraints are left out of the CS, and that the roster cost can be quantified.

Obviously, this cost includes "hard costs" such as the sum of the costs of the trips therein, but also can be designed to included the "soft costs" associated with deviations from crew preferences, by establishing penalties for the kind of, and extent of deviation from, each preference.

In the case of the PBS systems that we will be considering here, a unique set of parameters defining these penalties must be specified for each crew member. We will assume here that each roster in the CS is associated with one crew member, which eliminates the need to separately model crew-specific horizontal constraints such as fixed "carry-in" (i.e. the remainder of a trip carried over from the previous scheduling period), vacation, and training. We will also assume that each roster in the CS is "legal", i.e. satisfies all horizontal constraints, and has a unique cost associated with it, which includes the hard and soft costs discussed as well as preference weightings that are larger for more senior crew members.

The PBS (or Lines) problem can then be expressed as the following IP *set partitioning problem* (SPP):

minimize $C == \mathbf{cx}$ [1.a]
subject to: $\mathbf{Tx} = \mathbf{b} = \mathbf{1}$ [1.b.1]
$\mathbf{Hx} = \mathbf{b} = \mathbf{1}$ [1.b.2]
$x_i = \{1 \text{ or } 0\}$ [1.c.]

where C == total solution cost; $\mathbf{c} == \{c_i , i = 1 .. N_{lines}\}$ == row vector of line costs; $\mathbf{x} == \{x_i , i = 1 .. N_{lines}\}$ == column vector of line inclusion probability variables x_i, where x_i = {1 if line i is in the OLS, else 0}; $\mathbf{T} == \{T_{ji} , j = 1 .. N_{trips} , i = 1 .. N_{lines}\}$ == trip inclusion matrix with elements T_{ji}, where T_{ji} = {1 if trip j is in line i, else 0}; and $\mathbf{H} == \{H_{ki} , k=1 .. N_{lineholders} , i=1 .. N_{lines}\}$ == lineholder assignment matrix with elements H_{ki}, where H_{ki} = {1 if line i can be assigned to lineholder k, else 0}; and **1** == column vector each element of which equals 1. In order to prevent the same line from being assigned to more than one lineholder, matrix **H** is restricted to block diagonal structure, i.e. each line i is assignable to only one lineholder so that $(\mathbf{1}^T \mathbf{H})_i = 1$. This means that the same generic roster (i.e. sequence of trips) may be incorporated in several different lines, one for each lineholder that can fly it, but this is necessary because each different lineholder creates a different context for evaluating the cost and legality of the generic roster.

For the sequel, it will be convenient to combine **T** and **H** matrices into one constraint matrix **A == (T,H)**^{T} (i.e. with the rows of **T** followed by the rows of **H**), so that the IP SPP can be expressed:

minimize C == **cx** [1.a.]
subject to: **Ax = b = 1** [1.b.]
x_i = {1 or 0} [1.c.]

This Lines problem can be stated in English: Minimize solution cost subject to vertical constraints: (A.) Each trip is assigned to one and only one line. (B.) One and only one line is assigned to each line holder. There is thus one constraint for each trip and one constraint for each line holder.

Appendix 2: Solving the PBS Problem Using Column Generation (CG) Methods:

In this section we discuss how the PBS (or Lines) problem modeled above can be solved with column generation (CG) methods. (The major CG results discussed here can be understood in terms of a *Lagrangean relaxation* to the IP SPP problem stated above. The author has reviewed this Lagrangean relaxation, as applied to the airline crew pairing optimization problem ('Pairings'), in considerable detail in Tumpson, 2004. The reader is referred to this paper for detailed discussions of why column generation possesses the properties that it does.)

Column Generation denotes a class of techniques that start with a small CS of rosters and add a roster to the choice set only if its *reduced cost* (RC) is negative. Adding a roster k to the CS requires the addition of a "column" $\mathbf{A}_k == \{A_{jk}, j=1..N_{trips}+N_{lineholders}\}$ to the $\mathbf{A}$ matrix in constraints [1.b.], hence the term "column generation". To define the "reduced cost", and to motivate its use as a criterion for column generation, we begin with a slightly less constrained problem, the (primal) *LP set covering relaxation* of the Lines IP problem:

minimize solution cost $\mathbf{cx}$ [2.a.]
subject to set-covering constraints: $\mathbf{Ax} \geq \mathbf{b} = \mathbf{1}$ [2.b.]
$\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}$ [2.c]

and its corresponding dual problem:

maximize: $\mathbf{yb}$ [3.a.]
subject to constraints: $\mathbf{c} - \mathbf{yA} \geq \mathbf{0}$ [3.b.]
$\mathbf{y} \geq \mathbf{0}$ [3.c.]

Note that both the primal and dual problems have the same LP optimum = max($\mathbf{yb}$) = min($\mathbf{cx}$), and at optimum, $(\mathbf{c} - \mathbf{yA})\mathbf{x} = \mathbf{y}(\mathbf{b} - \mathbf{Ax}) = 0$. For a line i to be in the solution, x_i must be > 0 which means that *reduced cost* $\underline{c}_i(\mathbf{y}) == (\mathbf{c} - \mathbf{yA})_i$ must vanish. Alternatively, if $\underline{c}_i(\mathbf{y}) > 0$, as is allowed by dual constraint [3.b.], then x_i must be = 0, i.e. line i is not in the solution. This suggests that the value of $\underline{c}_i(\mathbf{y})$ can be used to decide whether a roster i should remain in the choice set: e.g., a positive value could be used as a criterion for eliminating roster i from the choice set.

Suppose we have two roster choice sets: a huge set CS_{max} of which a small set CS_0 is a subset. Suppose we solve the dual set covering problem on CS_0, obtaining dual solution $\mathbf{y}^0$. It can be shown (see Tumpson, 2004) that iff $\underline{c}_i(\mathbf{y}^0) \geq 0$ for all rosters i in CS_{max}, then dual solution $\mathbf{y}^0$ is optimal for CS_{max}. This has two important implications: (1.) the existence if a roster k in CS_{max} with $\underline{c}_k(\mathbf{y}^0) < 0$ means $\mathbf{y}^0$ is sub-optimal on CS_{max} and that adding roster k to CS_0 (to yield CS_1) will result in dual solution $\mathbf{y}^1$ with an improved optimum over $\mathbf{y}^0$. Thus, negative RC is a solution improving criterion for adding a roster (column) to the choice set. (2.) If no column with negative RC $\underline{c}_k(\mathbf{y}^0)$ can be found in CS_{max}, then $\mathbf{y}^0$ is optimal on CS_{max}. This is an amazing result in that optimality on a huge set can be proven by solving a much smaller set and scanning the larger set and thereby finding no negative RC columns.

Define a "complete" subset CS_c of CS_{max} to be any subset of CS_{max} such that if $\mathbf{y}^c$ is an optimal dual solution to CS_c, then $\underline{c}_i(\mathbf{y}^c) \geq 0$ for all trips i in CS_{max}. Thus any complete subset of CS_{max} fulfills the necessary and sufficient conditions for its dual solution to be optimal also for CS_{max}.

A simple column generation procedure for creating a complete subset of a very large CS_{max}, then, is as follows:

Initialize: n to 0, CS_n to initial CS;
(A): solve CS_n for dual solution $\mathbf{y}^n$ and find a subset S_n of columns k in CS_{max} such that $\underline{c}_k(\mathbf{y}^n) < 0$;
if set S_n is empty: then CS_n is complete: DONE;
else: set CS_{n+1} to $CS_n \cup S_n$; increment n; go to **(A):**

Strengths and Weaknesses of Column Generation:

Column Generation arose in response to the problem of applying LP methods to huge numbers of variables. In the case of both crew pairings and lines optimization, the choice sets of scheduling objects (trips and rosters respectively) are huge: potentially on the order of billions. (See estimate of PBS problem size.) In both cases, the scheduling object is the smallest to which legality can be ascribed and a cost can be associated, which are minimal conditions for an optimal solution to be extracted from a choice set with LP methods. The choice sets in both cases are way too large to be subjected to standard LP methods such as simplex. If one attempts to reduce the size of the choice set in some intelligent manner to some small and manageable subset, one runs the risk of throwing the baby out with bath water: unless that subset contains an optimal solution to the larger choice set, its solution will obviously be suboptimal.

As described above, CG provides a criterion (negative RC) for including new columns into an evolving CS, columns which will improve the optimal solution that may be extracted from each successive CS. But there are two potential problems. First, since the problem being solved is an LP relaxation of the original IP problem, the LP optimum may not correspond to the IP optimum, nor is there any quick and easy way to extract the best IP solution from the LP solution. (This problem can be addressed by using the IP method of Wedelin, 1995, in that solutions are innately IP. This method has the innate flaw that it may fail to converge to an IP optimum.) Second, there is no straightforward way to decide which is the most efficient order to add negative RC columns to the choice set; obviously if the wrong ones are added, there will be little or no improvement, and this could occur over and over until the CS is too big to solve with LP methods.

One major approach to ordering the column generation is to build evolving tree graphs whose nodes and edges correspond to components of the scheduling objects (e.g. legs and ground connections are the node and edge components of trips, and activities and days off are the node and edge components of rosters). Branch and price techniques require that the order in which these trees are built corresponds to the accumulating RC of the components; when a scheduling object is completed with the most negative RC, it is added to the CS. This seems like a reasonable criterion for adding columns, but there is no assurance that such ordering will actually lead most directly to an optimum solution.

In the case of Pairings, very good results have been obtained. (See e.g. Anbil, Forrest, and Pulleyblank, 1998, Vance *et al*., 1997, and Wedelin, 1995.) But the Lines problem differs from the Pairings problem in that there are many more combinations of rosters which are of comparable quality, so adding such columns to a choice set may not appreciably improve the objective but might instead diminish the integrality of the solution as the overlapping rosters of similar quality yield solution inclusion probabilities having similar non-integral magnitude.
To deal with this Lines specific tendency for CG choice sets to not converge toward small subsets containing optimal IP solutions, providers have tried heuristics for column generation other than RC-based branch and price, such as artificially restricting the solution space (e.g. by fixing certain trips in certain solution lines) and generating columns which are various combinations of the freed components with the fixed ones. (This approach is described in Kohl and Karisch, 2004.) This, of course, amounts to a local search of solution space in the vicinity of the fixed component structures, and while it may yield improvements, it still relies on heuristics to determine, somewhat arbitrarily, which local search to pursue.

This suggests that for Lines, unlike for Pairings, CG guided by RC and branch and price scanning heuristics may not converge quickly to a provably near-optimum result. The fault, of course, may ultimately lie in the reliance on LP techniques that require such complex and numerous scheduling objects. It may be that for Lines, heuristics which combine the simpler component scheduling objects, i.e. trips, may arrive more quickly at very good if not near optimal results. If such heuristics can indeed quickly generate good solutions, this makes them a viable alternative to CG and other LP methods, especially in circumstances, such as schedule recovery, which require speedy adjustment of schedules.

Appendix 3: Incompatibility of Non-Interactive Bidding and Parameter Optimization:

Define a *preference parameter* to be a crewmember-specific penalty applied to a trip's rating per unit deviation from that particular preference characteristic. E.g. if a crew member prefers trips that depart at a particular time, an appropriate preference parameter would be "units of penalty per hour deviation from that departure time". Obviously if a crew member has no preference with regard to a characteristic, the preference parameter has value zero.

A non-interactive system uses the one set of preference parameters that are submitted by the crewmembers in their one (blind) bid. As mentioned above, these blind bids can lead to clashes between crew members in their preferences which limit the quality of the solution.

A non-interactive system could attempt to reconcile preference clashes by optimizing preference parameters, i.e. by not fixing each crewmember's preference parameters to the values specified in his bid, but instead allowing them to have values within an allowable range around the bid values, and then finding parameter values within those ranges which lead to an optimal solution when referred to, say, the originally bid parameter values.

Parameter optimization is a natural and potentially useful response to optimization problems where (1.) the parameters are "soft", i.e. they are intended to quantify a characteristic that is to a large extent subjective in nature, and (2.) the space of solutions and their costs for any single set of parameters is characterized by many "local minima" (i.e. the solution cost is locally minimum and increases locally in all directions) toward which the solution process might be drawn, but which are not in fact the global cost minimum sought by the optimization process.

The need for "soft" parameters is obvious: adjusting parameters is necessary for the process to work, and cannot be applied to "hard" parameters which cannot be adjusted. The whole point of parameter optimization is in fact to change the topography of the solution space so that when the process gets stuck in a local minimum, the cost topography can be changed (by changing preference parameters) so that a local minimum of the old cost space is transformed into a slope in the new space down which a local search may proceed toward the minimum.

The PBS problem partakes of these two characteristics, and so is a candidate for parameter optimization. But such a parameter optimization approach when applied to a non-interactive PBS has the disadvantage that, since adjustments to crew parameters are not ratified by the crew, there is no guarantee that the resulting solution will even reflect the actual preferences of the crew anymore.

Appendix 4: Estimate of the PBS Problem Size

As an example of the crew rostering problem, consider a Continental Airlines schedule with about $N_t = 1100$ trips to assign to $N_{lh} = 275$ line holders, for an average of $M = N_t / N_{lh} = 4$ trips per line holder. Assume that a line can be partitioned into M time sequential sectors, and that the N_t trips can be partitioned into M equal groups each of which fits into one of the M sectors. The total number of *anonymous* rosters which can be built (without violating trip overlap constraints) is then approximately $(N_t / M)^M = N_{lh}^M = 275^4 = 5.72$ billion.

Anonymous rosters are rosters which have no fixed activity elements (such as carry-in, vacation, training) which would particularize them to a particular line holder. Thus we would have to map each anonymous roster to each of the N_{lh} trip-empty line holder rosters, to create a maximum of $N_{lh}^{M+1} = 275^5 = 1.57$ trillion rosters in the full roster CS. Of course, many of these mappings are illegal in that one or more of the trips overlaps the fixed activity elements, but even if only 1 in 300 of the anonymous rosters legally maps onto each of the N_{lh} trip-empty line holder rosters, then there are still over 5 billion rosters in the full roster choice set CS_{max}.

References

Anbil, R., Forrest, J., and Pulleyblank, W., "Column Generation and the Airline Crew Pairing Problem", Documenta Mathematica, Extra Volume ICM 1998, III, 677--686

Kohl, N., Karisch, S., "Airline Crew Rostering: Problem Types, Modeling, and Optimization", Annals of Operations Research 127, 223–257, 2004

Sellmann, M., Zervoudakis,K., Stamatopoulis, P., and Fahle, T., "Crew Assignment via Constraint Programming: Integrating Column Generation and Heuristic Tree Search", Annals of Operations Research 115, 207–225, 2002

Tumpson, D. "Solving the Crew Pairing Problem with Diad and Trip Choice Sets -- I. Trip-based Methods involving Reduced Cost and Column Generation", 2004

Vance, PH, Atamturk, A, Barnhart, C, Gelman, E, Johnson, EL, Krishna, A, Mahidhara, D, Nemhauser, GL, Rebello, R, "A Heuristic Branch-and-Price Approach for the Airline Crew Pairing Problem", 1997

Wedelin, D. "An algorithm for large scale 01 integer programming with application to airline crew scheduling", Annals of Operations Research, vol. 57, pp. 283301, 1995