Some Results on Guidelines for Treatment of Traffic Congestion on Street Networks

William R. McShane and Louis J. Pignataro, Department of Transportation Planning and Engineering, Polytechnic Institute of New York, Brooklyn

The guidelines reported in this paper are intended for use by the practicing traffic engineer. The concept of the guidelines, the recommended approach, and a framework for addressing a particular problem are outlined. The framework includes both preliminary identification of the cause of the problem and categories of the treatments available. A sequence of treatment classes is recommended. Proper signalization and provision of added space (bays, lanes, and the like) are of prime importance. Some of the important results or recommended actions are presented in individual sections within the paper—shorter cycle length to avoid spillback, equity offsets to compensate when oversaturation occurs, and a number of nonsignal considerations. Some of the flow charts to aid in decision making are shown.

The problem of traffic congestion, traffic saturation, and traffic oversaturation presents traffic engineers with one of their most difficult tasks. NCHRP Project 3-18(2) (1) addressed this topic. As part of that project, a document on guidelines was prepared. This paper reports on the structure of those guidelines and on some relevant specific results contained therein.

The guidelines are intended to aid traffic engineers in executing their duty by reminding them of available options, by uncovering some subtleties that can be overlooked, and by presenting quantitative insight into the relative benefits of various options. In this way, an appreciation can be obtained of when various options are effective or necessitated or both, of how much impact can be expected, and what combinations are most effective.

SUMMARY OF THE GUIDELINES

Unequivocal statements of when particular techniques or combinations of techniques are better than others could not be developed. However, certain categorical statements can be made, and a logical analysis framework can be specified.

There is a logical set of steps to take to treat the problem of congestion and saturation.

1. Address the root causes of congestion first, foremost, and continually.
2. Update and, if necessary, improve the signalization.
3. Provide more space by use of turn bays and parking restrictions.
4. Consider both prohibitions and enforcement realistically to determine whether an effort would be futile or whether it might merely transfer the problem.
5. Take other available steps, such as allowing right turn on red (RTOR) while recognizing that the benefits will generally not be so significant as either signalization or more space.
6. Develop site-specific evaluations where there are conflicting goals, such as providing local parking versus moving traffic.

The following sections provide an exposition of these key elements in the recommended method of approaching the problem. The framework by which a problem should be considered is then presented. The framework has two components:

1. A focus on the identification of the problem in terms of probable cause and
2. A focus on the categorization of the possible solutions so that they may be readily found within these guidelines.

Root Causes

The problem should be attacked at the root causes, which are

1. Land use policies (concentrations of movement implied in some land use distribution, use of on-street facilities for loading and unloading of goods, and multiplicity of access and egress points and standing queues on the rights-of-way);
2. Demand pattern (concentration of work trips in a short period and unrestricted hours of goods activity);
3. Size of demand (number of vehicles, particularly low-occupancy private automobiles);
4. Use of street space [inefficient curb space management (parking, moving lanes, and the like)]; and
5. Pedestrian conflicts (lack of grade separation in areas of extreme intensity).

Engineers should continually educate other specialists and the public about this need to attack root causes. However, in the time frame of local, site-specific problems that they must address, the guidelines must often suffice.

Signalization

It is difficult to overstate how often poor signalization is the basic problem. After the signalization is improved through reasonably short cycle lengths, proper offsets (including queue clearance), and proper splits, many problems disappear. Sometimes, of course, there is just too much traffic. At such times, equity offsets to aid cross flows and different splits to manage the spread of congestion are appropriate if other options cannot be called on.

Discussions with and surveys of traffic engineers have revealed that a systematic consideration of signalization for congestion and saturation is rarely done. The procedures contained in the guidelines are recommended for use. Study of representative traffic patterns lends strong credibility to the conclusion that minimal-response (preplanned) signals policies generally suffice.

Space

If a problem cannot be remedied by signalization, then more space may be needed. Left-turn bays and, where appropriate, right-turn bays can aid individual movements as well as remove impediments to the through flows. Without question, additional lanes are a benefit. However, this tends to be an arterial-long solution that engineers often don't like.

Two-way left-turn lanes offer special advantages particularly along strip development sites. One-way systems, arterials with unbalanced lanes, and reversible lanes offer advantages, but also represent either major implementation problems or site-specific treatments. One-way systems require studies quite beyond the scope of congestion, although that may be the prime motivator for such a study. Unbalanced lanes require certain volume patterns to be useful.

Prohibition and Enforcement

Before instituting any prohibition or enforcement program, the engineer must decide whether it can be enforced strictly enough to realize most or all of the projected benefit (curb parking prohibition to provide a moving lane) and whether it will simply transfer or even accentuate the overall problem (circulation of vehicles that would otherwise be double-parking). Only then can the engineer consider that there is a potential benefit.

Other Attempts

Some solutions that are available can have either a net benefit or a net disbenefit depending on the site and the situation. RTOR is such a case. If it allows vehicles to "escape" from a congested arterial, it is quite suitable. If, however, it allows vehicles to "steal" available space on such an arterial, then it is inappropriate.

The question of prohibitions such as those affecting turning arises. These can be used only if alternate routes exist. Often, this is not the case.

More Detailed Evaluation

Very often, application of these guidelines will clarify the issue and identify a solution. In some cases, the final decision will rest on conflicting desires that might be usefully viewed in economic terms. Is removal of five parking spots worth the delay savings to the traffic stream? Are off-street goods facilities justified economically? Are pedestrian phases justified in terms of total person-minutes saved? What is a proper allocation of curb space?

If necessary, engineers can develop such an analysis for their individual cases. More general treatment of such situations is recommended for future research. Some of this type of work on curb space management for goods facilities has been done (2).

RANGE OF SOLUTIONS AVAILABLE

The fact that there is substantial traffic congestion virtually ensures that one is dealing with signalized intersections. However, it does not follow that one has only signal remedies at hand. Indeed, the possible treatments may be broadly classified as signal (timing and coordination) and nonsignal treatments.

Within the signal classification, there are two major subclassifications: minimal response (preplanned) and responsive signal control. It is not at all well established that highly responsive control is better than preplanned signal plans particularly for the heavier flow range. This is an indication that is being reinforced by trends in major computer-based study projects. Within the nonsignal classification, there are also two major subclassifications: regulatory and operations. Regulatory action consists of enforcement and of prohibitions. Operations, as used herein, consists of all other traffic measures.

The role of enforcement cannot be minimized. Many problems can be traced to the lack of enforcement of existing traffic regulations. In other cases, certain treatments are precluded initially because it is anticipated that there will not be adequate enforcement to have the measure work. Within this section, the following topics are addressed:

1. Improvement to be sought and
2. Available solutions.

The material on these topics combined with the method of approaching the problem previously presented represents the essence of the recommended framework for attacking the problem of congestion and saturation.

Improvements Desired

Before enumerating possible treatments, it is appropriate to dwell on the ends to be achieved. In other words, what improvements are being sought? The following is a set of the most common improvements that traffic engineers may wish to make when they face a traffic congestion problem:

1. Reduced geographic spread of congestion,
2. Reduced rate of spread of congestion,
3. Increased throughput,
4. Reduced delay,
5. Reduced stops, and
6. Improved regularity of service.

These are stated in the broad terms usually encountered as goals or objectives. Some of the items in this list are really secondary to other items for given flow levels. Figure 1 shows the primary objectives that should be sought by the engineer. These are dependent on the traffic level. First and foremost, the engineer must realize that at the more extreme flow levels the objective be-

comes the avoidance of spillback. All else follows from this. This is the explicit objective. The mathematical niceties of minimum stops or minimum delay or both collapse in the face of intersections blocked by vehicles.

Some comments on Figure 1 are in order. First, the primary objective to which engineers should address themselves does change depending on the flow level. Second, the major index of performance [measure of effectiveness (MOE)] also changes. However, both sets are well correlated to queue-extent measures; therefore, queue or occupancy patterns or both—particularly during red and at the onset of green—are good indicators across the entire range of conditions.

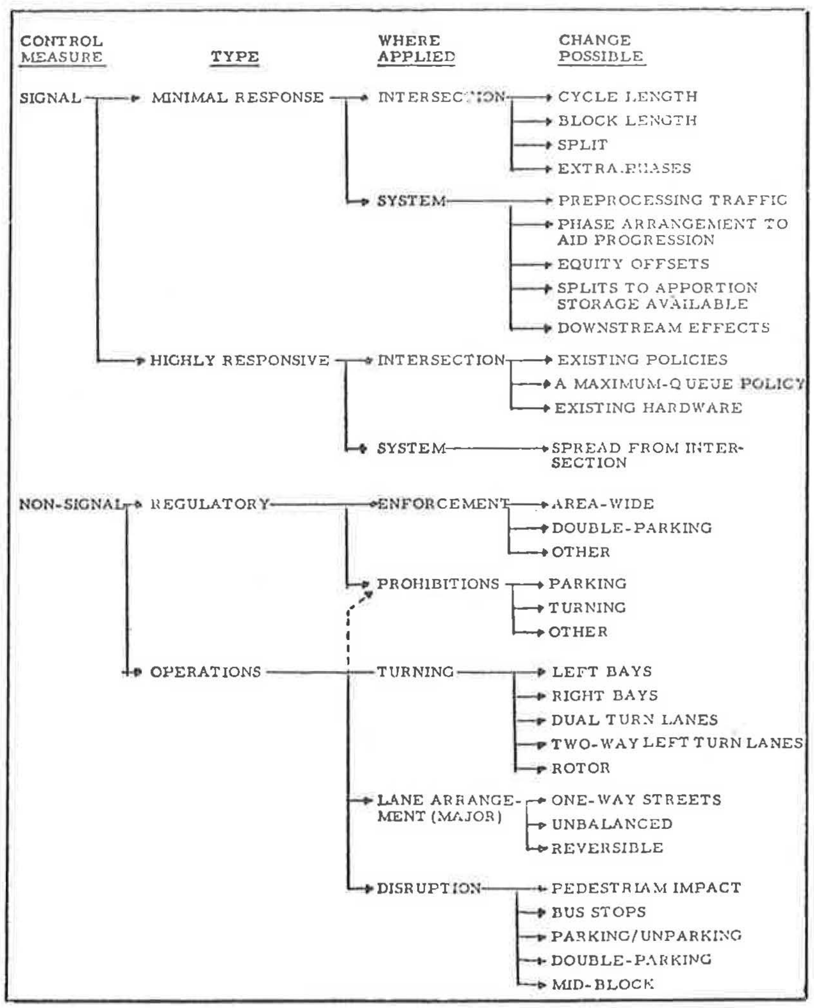
A special word of attention is appropriate on the desire to improve the regularity of service. During simple congestion, the variance as well as the mean of the delay per cycle increases as demand approaches capacity. Thus the individual driver will be exposed to greater variability in his or her individual experience from day to day. Improvements that minimize the mean delay will also enhance the regularity of the delay suffered.

There are times when a basic solution has been

Figure 1. Dependence of desired objectives on traffic condition.

TRAFFIC CONDITION	Congestion →	Saturation/Oversaturation
OBJECTIVE TO BE DESIRED	{Minimize Delay, Minimize Stops} →	{Avoid Spillback, Provide Equitable Service}
IMPACT OF USING OR TRYING TO ACHIEVE ABOVE OBJECTIVE (i. e., RESULTS)	• Intersection-specific congestion, not area-wide • Throughput Adequate • Service Regularity Enhanced • Enhancements to Special Groups (such as right-turners) Possible	• Reduce Spatial Extent and Rate of Spread of Congestion • Potential for occurrence of area-wide Congestion reduced, and service regularity thus enhanced • Throughput not Impeded • Special Treatment of User Groups Possible
MAJOR INDEX OF PERFORMANCE	Delay, Stops	Queue Extent and Occupancy

Figure 2. Range of solutions available.



achieved and some possible benefits can be realized by additional improvements addressed to specific subgroups. Right-turn bays are examples, for they frequently have little impact on a measure such as average delay of all vehicles although they have truly substantial benefits for a smaller segment of the traffic stream—those turning right.

Solutions Available

Figure 2 shows the range of solutions available. There is no simple statement of an ordered list of recommended solutions in decreasing order of preference. There are, however, indications of how much of one solution must be implemented to have an equivalent impact of so much of another solution. The engineer must then use this knowledge in conjunction with local conditions and practices to reach a decision. There are also indications of how best to use two solutions in conjunction with each other.

It can be stated that there is a simple set of initial steps that can be followed as an elimination checklist (Figure 3).

The engineer must reach a preliminary judgment of the underlying cause of the problem. At the same time, he or she must be assured that the solution is not trivial. Extensive queues may drive an engineer into the depths of these guidelines too quickly. Such problems can arise because of poor offsets, outdated splits, and excessive cycle lengths. As a first step, therefore, the engineer should prepare a preliminary opinion on the underlying cause. Given a preliminary opinion, there are a number of possible solutions that one is tempted to consider. Much of the guidelines are addressed to the candidate solutions, the considerations involved, and the relative merits of each.

CYCLE LENGTH AND BLOCK LENGTH

Two questions must be addressed. Do long cycle lengths have any virtue in their own right? Does block length enter into the cycle length determination?

Cycle Length and Capacity

One of the most prevalent erroneous beliefs in the traffic engineering community is that the capacity of an intersection increases substantially as cycle length is increased. This has been questioned in the past (3), and studies (4, 5) have provided data to support such questioning. Lack of substantial capacity increases with increasing cycle length is rooted in at least three factors: (a) Loss time per cycle is not that severe because of both usage of the amber and lower start-up delays than are often assumed; (b) the use of longer greens is inefficient because of increasing headways; and (c) a demand to fill rather long green times cannot be provided. The last item is just another manifestation of the storage problem.

Block Length and Storage

Cycle length may not be as powerful a capacity improver as one might think. However, no evidence was offered that there is a positive good to short cycle lengths in some cases.

To avoid a high potential for spillback, a minimum condition is that the moving platoon not exceed the available link storage. Thus

$$\pounds \leqslant f_1(\xi C/3600) \qquad (1)$$

where

ξ = vehicle storage length,
$\pounds$ = link storage distance,
C = cycle length in seconds, and
f_1 = flow in passenger cars per hour per lane.

$\pounds$ need not be the physical length of the link. If a policy decision is made that the stored vehicles should come no closer than within 61 m (200 ft) of the upstream intersection, then $\pounds$ is 61 m (200 ft) less than the physical length. Such a policy decision is in accord with the avoidance of the perception of congestion.

In order to avoid the situation shown in Figure 4, excessively long platoons must be avoided. Equation 1 may be rewritten as a constraint on cycle length:

$$C \leqslant (3600/f_1)(\pounds/\xi) \qquad (2)$$

Clearly, two contrary forces are at work. As the total critical land flow (all approaches) increases, cycle length increase brings some benefit; at the same time, flow increases on any one approach decrease the maximum cycle length permitted. Figure 5 shows block length that is adequate for cycle length set.

EQUITY OFFSETS

Unfortunately, avoiding spillback is not always possible, for there may be too many vehicles attempting to enter a particular link. With the extreme traffic congestion, it is not uncommon to see vehicles storing themselves in the intersection, to the detriment of the cross traffic. One common solution to this spillback problem is to place a traffic control officer at this site to prevent such events. Another approach is an intensive ticketing program for such offenders. The former approach is not only historically more effective, but it is also the one demanded by a public afflicted with spillback.

A possible alternative solution exists in changing the basic concept of what the offset is supposed to accomplish. However, this solution should not be implemented until one is certain that a better offset cannot alleviate the problem. The treatment to be presented now is only for that period after the best possible offset has failed because of the size of the volume demanding access to the link.

The treatment, shown in Figure 6, is as follows:

1. Allow the oversaturated direction to have green until the vehicles blocking the intersection just begin to move;
2. Switch green to the cross traffic; and
3. Allow the cross stream to move until it has had an equitable input into the oversaturated link or at least to the intersection.

This offset, the equity offset, is not determined in the usual fashion. The upstream red should begin L/V_{ACC} s after downstream green initiation, where V_{ACC} is the acceleration wave speed in meters per second. Assume g_1 as the green time at the upstream intersection (percentage of cycle) and g_{c1} as the green time at the critical intersection. Thus

$$t_{off} = g_1 C - (L/V_{ACC}) \qquad (3)$$

where C = cycle length in seconds and L = physical block length in meters. Typically, $V_{ACC} \simeq 5$ m/s (16 ft/s).

The original link must have unavoidable saturation. Neither any signal nor any available nonsignal remedy could have helped it. Only then is this link given up on and the best possible done for other traffic.

Figure 7 shows an arterial on which the volume en-

sures oversaturation, at least of links 2 and 3. There are no turns. The equity offset for link 2 is computed as -1.5 s so that a simultaneous offset will happen to provide an equity offset function.

Figure 8 shows the queue per lane on link 1, the cross-stream link. The only offset that is varied is that in link 2. Figure 8 clearly indicates that the equity offset (offset = 0) is quite important to the cross-street traffic (link 1).

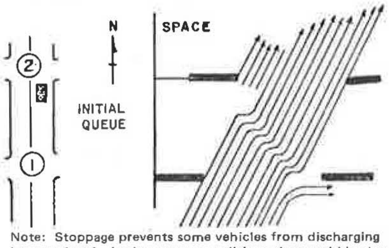
DOWNSTREAM BLOCK LENGTH

If a decision is made that the cycle length at the critical intersection (CI) is to be larger than the downstream

Figure 3. Initial classification and elimination checklist.

Apparent Problem	Initial Steps
Area-Wide Congestion	• Identify the Critical Intersection (CI). • It is unlikely that there are two CI's. If there are, it is likely that each can be considered with its own area of influence. • Do not erroneously identify the intersection downstream of the CI as the CI. • Classify the type of oversaturation
Spillback in a Link	• Determine whether a simple split adjustment is sufficient. • Determine whether the cycle length is too long for the block length and flow. • Determine whether the offset is poor for the primary and secondary flow mix. • Identify any special blockages in the link (double parking, queues for garages, car washes, etc.).
Single Intersection	• Isolate primary symptom
→ Single Approach	• Check split and offset as above. • Check burden caused by turns.
→ Two Approaches, same Right-of-Way	• Check same as one approach, but with emphasis on interference with each other.
→ Two or More Approaches More than One Right-of-Way	• Consider methods for increasing net capacity. • Consider methods for minimizing spatial extent of possible oversaturation and area-wide congestion.

Figure 4. Cycle length too large for block length and flow.



Note: Stoppage prevents some vehicles from discharging intersection 1. It also causes conditions that could lead to a blockage of intersection 1. For example, if there were any delays in moving through link 1-2, the cross stream traffic at intersection 1 would be quickly affected.

Figure 5. Block length adequate for cycle length set.

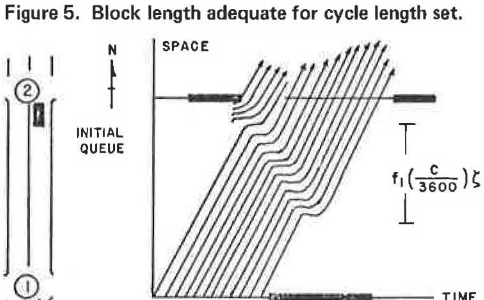


Figure 6. Equity offsets.

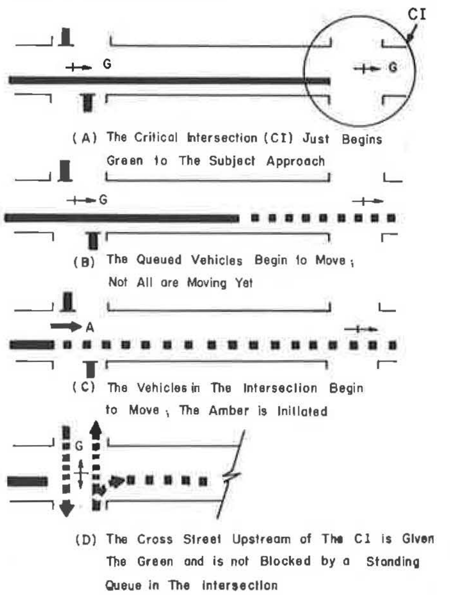


Figure 7. Impact of offsets.

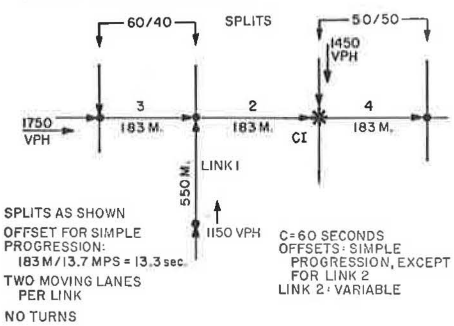
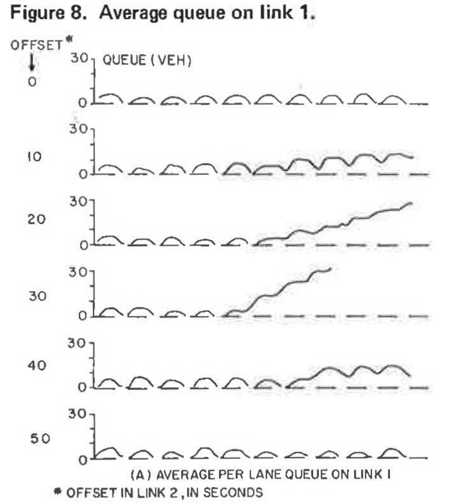


Figure 8. Average queue on link 1.

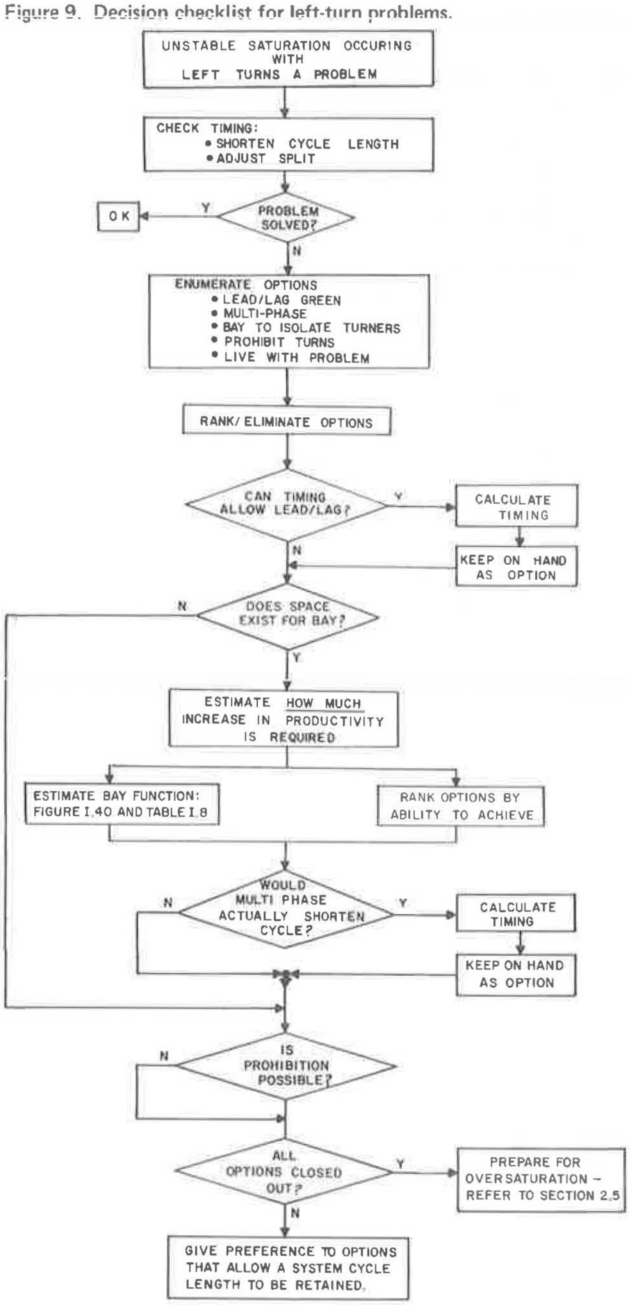


cycle lengths (perhaps because the CI is to have multiple phases), the queue extent to be stored may be shown to reach a maximum of more than twice the single-cycle discharge of the CI into the link.

SPLIT

For congested flow, the standard rule of proportioning available effective green in the ratio of the critical lane flows will not suffice. It should be appreciated that, as the demand approaches capacity, greater queues will be experienced, as will greater delays and greater fluctuations (variance) in delay per cycle. For unstable saturation and for oversaturation, a different concept should prevail. Clearly, there are situations in which the CI simply cannot handle the total demand put on it. Must the same sense of equitable treatment still hold? We recommend that the split be apportioned so that the rate of growth of congestion in both (or all) directions be equalized; both directions should exceed their respective links or defined system boundaries at the same time. This is addressed in the guidelines.

Figure 9. Decision checklist for left-turn problems.



EXTRA PHASES

As a rule, multiple phases should be avoided particularly because they generally require an increase in the overall cycle length. Other options should be considered: turn bays, shorter cycle lengths, parking restrictions, leading or lagging greens or both, and turn prohibitions. Still, there are cases when multiple phasing is clearly necessary. Even when the left-turn volumes are less than 120 vehicles/h, there are conditions under which a left-turn phase can be added without increasing the cycle length required.

ENFORCEMENT AND IMPACT

Two of the most chronic violations that aggravate the congestion and oversaturation problems are intersection blockage and parking regulation violations. Equity offset represents an attempt to circumvent the first and avoid or delay the need for on-scene traffic control officers.

The UTCS-1 simulation was used to study the impact of double-parkers in a 183-m (600-ft), 3-lane link. The resultant curve can also be used to estimate the impact of adverse uses of a curb lane from which parking was removed to increase capacity.

RIGHT-TURN BAYS

The creation of a right-turn bay allows

1. An increase in productivity or, if it is desired, a decrease in the effective green allocated to the phase and
2. A decrease in the local delay with the right turners realizing most of the delay savings.

The increase in productivity, expressed as a percentage, can be comparable to the right-turn percentage.

The length of the turn bay should be approximately the same length (slightly longer) as the queue that typically forms. In this way, maximum presorting can occur. Thus, in practical terms, short cycle lengths and this objective complement each other, for released platoons are smaller, and the necessary length is easier to achieve.

An example in the guidelines illustrates, however, that much of the benefit is achieved by the existence of a bay of even moderate size. Still, short cycle lengths aid presorting and should be used as a companion measure.

PROBLEMS DUE TO LEFT-TURN MOVEMENT

Figure 9 shows a decision checklist to be used in considering a problem that arises because of a left-turn movement. The final decision must be evaluated with due consideration of the three problems.

1. Is an alternate route available for the left turners? How much does it adversely impact them? Can the alternate route afford to be impacted by the additional flow?
2. How many parking spaces must be removed to aid the flow means of a turn bay or even an additional lane? What is their economic value?
3. What delay is being suffered now?

Frequently, the decisions can be reached by systematically thinking of the options as shown in Figure 9 and keeping these issues in mind. Sometimes a benefit-cost or cost-utility decision would be required for a "close de-

cision" or highly sensitive issue.

TWO-WAY TURN LANES VERSUS UNBALANCED FLOW

The engineer may judge that heavy congestion or even stable saturation at intersections is sometimes inevitable. At mid block, however, it is the opinion of some that the engineer, and the public, will generally find a lower (but significant) amount of congestion to be equally unattractive. The engineer therefore may be solving a congestion problem at mid block and a saturation problem at the intersection.

Two-way left-turn lanes can remedy such congestion impact if space permits. There is some evidence that they can substantially improve the accident situation (6). The option of a two-way turn lane may solve the mid-block congestion problem. Given that the additional lane will now be added, however, opens the possibility that it can be used at the intersection by the through flow—if unbalanced flow is implemented. This will allow a reduced green for this approach, perhaps to the benefit of other phases and thus the system. The guidelines contain a checklist similar in concept to the information shown in Figure 9 for mid-block congestion or for unbalanced flow and reversible lanes. Of course, this use of unbalanced flow requires planning. Can the opposing direction accommodate its own turners without unduly impeding its continuing vehicles?

Note that any decision involving two-way turn lanes versus unbalanced flow in this context considers only the congestion and saturation issue. Data on accident advantages are not sufficient to say whether there is an accident benefit of two-way turn lanes (for example, removal from the traffic stream as opposed to decreasing the density) that should override.

DUAL TURN LANES

When turning volumes are extremely heavy, both capacity and queue extent may dictate use of two lanes for a turning movement. Data in the literature and discussions with engineers responsible for such sites indicate that there is no downward correction factor needed for either of the lanes.

OTHER NONSIGNAL REMEDIES

The guidelines incorporate consideration of all remedies (possible treatments) shown in Figure 2 including left-turn bays and RTOR. Other results are incorporated into the guidelines, including

1. Rules of thumb on when to use simultaneous and other progressions,
2. Rules of thumb for productivity increases due to left-turn bays,
3. Circumstances under which multiple phases may actually not increase the cycle length, and
4. Illustrations of the relative impact of alternative treatments.

More important than any of these specifics, however, is the tutorial approach of the guidelines and the development of a systematic methodology. Once acquainted with these guidelines, engineers can better sort out the issues in their own applications and consider more implications.

CONCLUSIONS

The problem of congestion and saturation is widespread, and is not approached in any consistent manner. There are definite measures that can be taken, but preventive action addressing the root causes must be given a high priority. Among the measures that can be taken, those relating to signalization generally can have the greatest impact. There are distinct signal plans for avoiding spillback and for living with spillback. The nonsignal remedies are in no way to be minimized, particularly those that provide space either for direct productivity increases or for removing impedances to the principal flow. The guidelines produced in this work provide both a tutorial and an illustrated reference in what techniques to consider and how to consider them systematically. The interested reader is referred to Appendix 1 of NCHRP Project 3-18(2) (1) for a more detailed exposition.

ACKNOWLEDGMENT

This study was conducted under National Cooperative Highway Research Program Project 3-18(2) (1). The opinions and findings expressed or implied in this paper are ours. They are not necessarily those of the Transportation Research Board, the National Academy of Sciences, the Federal Highway Administration, the American Association of State Highway and Transportation Officials, or the individual states participating in the National Cooperative Highway Research Program.

REFERENCES

1. Traffic Control in Oversaturated Street Networks. NCHRP, Project 3-18(12), final rept.
2. P. A. Habib. Accommodating Goods-Movement Vehicles in the City Center. Polytechnic Institute of New York, Brooklyn, PhD dissertation, June 1975.
3. K. Moskowitz. Long Cycles, Short Cycles, and "Lost Time" at Signalized Intersections. Paper presented at the Workshop on Intersection Capacity, TRB, Jan. 1974.
4. G. M. Webb and K. Moskowitz. Intersection Capacity. Traffic Engineering, Jan. 1956.
5. R. L. Carstens. Some Traffic Parameters at Signalized Intersections. Traffic Engineering, Aug. 1971.
6. M. R. Hoffman. Two-Way Left-Turn Lanes Work. Traffic Engineering, Aug. 1974.