Signal Timing Manual - Second Edition (2015)

DETAILS

322 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-30888-5 | DOI 10.17226/22097

CONTRIBUTORS

Urbanik, Tom, Tanaka, Alison, Lozner, Bailey, Lindstrom, Eric, Lee, Kevin, Quayle, Shaun, Beaird, Scott, Tsoi, Shing, Ryus, Paul, Gettman, Doug, Sunkari, Srinivasa, Balke, Kevin, and Darcy Bullock; National Cooperative Highway Research Program; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

SUGGESTED CITATION

National Academies of Sciences, Engineering, and Medicine. 2015. *Signal Timing Manual - Second Edition*. Washington, DC: The National Academies Press. https://doi.org/10.17226/22097.

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

NCHRP REPORT 812

Signal Timing Manual

Second Edition

Tom Urbanik
Alison Tanaka
Bailey Lozner
Eric Lindstrom
Kevin Lee
Shaun Quayle
Scott Beaird
Shing Tsoi
Paul Ryus
KITTELSON & ASSOCIATES, INC.
Portland, OR

Doug Gettman
KIMLEY-HORN AND ASSOCIATES, INC.
Phoenix, AZ

Srinivasa Sunkari
Kevin Balke
TEXAS A&M TRANSPORTATION INSTITUTE
College Station, TX

Darcy Bullock
PURDUE UNIVERSITY
West Lafayette, IN

Subscriber Categories
Operations and Traffic Management

Research sponsored by the American Association of State Highway and Transportation Officials
in cooperation with the Federal Highway Administration

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2015
www.TRB.org

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

Systematic, well-designed research is the most effective way to solve many problems facing highway administrators and engineers. Often, highway problems are of local interest and can best be studied by highway departments individually or in cooperation with their state universities and others. However, the accelerating growth of highway transportation results in increasingly complex problems of wide interest to highway authorities. These problems are best studied through a coordinated program of cooperative research.

Recognizing this need, the leadership of the American Association of State Highway and Transportation Officials (AASHTO) in 1962 initiated an objective national highway research program using modern scientific techniques—the National Cooperative Highway Research Program (NCHRP). NCHRP is supported on a continuing basis by funds from participating member states of AASHTO and receives the full cooperation and support of the Federal Highway Administration, United States Department of Transportation.

The Transportation Research Board (TRB) of the National Academies of Sciences, Engineering, and Medicine was requested by AASHTO to administer the research program because of TRB's recognized objectivity and understanding of modern research practices. TRB is uniquely suited for this purpose for many reasons: TRB maintains an extensive committee structure from which authorities on any highway transportation subject may be drawn; TRB possesses avenues of communications and cooperation with federal, state, and local governmental agencies, universities, and industry; TRB's relationship to the Academies is an insurance of objectivity; and TRB maintains a full-time staff of specialists in highway transportation matters to bring the findings of research directly to those in a position to use them.

The program is developed on the basis of research needs identified by chief administrators and other staff of the highway and transportation departments and by committees of AASHTO. Topics of the highest merit are selected by the AASHTO Standing Committee on Research (SCOR), and each year SCOR's recommendations are proposed to the AASHTO Board of Directors and the Academies. Research projects to address these topics are defined by NCHRP, and qualified research agencies are selected from submitted proposals. Administration and surveillance of research contracts are the responsibilities of the Academies and TRB.

The needs for highway research are many, and NCHRP can make significant contributions to solving highway transportation problems of mutual concern to many responsible groups. The program, however, is intended to complement, rather than to substitute for or duplicate, other highway research programs.

NCHRP REPORT 812

Project 03-103
ISSN 0077-5614
ISBN 978-0-309-30888-5
Library of Congress Control Number 2015946872

NOTICE

The report was reviewed by the technical panel and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the National Academies of Sciences, Engineering, and Medicine.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; or the program sponsors.

The Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; and the sponsors of the National Cooperative Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

Published reports of the

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

are available from

Transportation Research Board
Business Office
500 Fifth Street, NW
Washington, DC 20001

and can be ordered through the Internet by going to

http://www.national-academies.org

and then searching for TRB

Printed in the United States of America

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Ralph J. Cicerone is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org**.

The **Transportation Research Board** is one of seven major programs of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to increase the benefits that transportation contributes to society by providing leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation.

Learn more about the Transportation Research Board at **www.TRB.org**.

COOPERATIVE RESEARCH PROGRAMS

CRP STAFF FOR NCHRP REPORT 812

Christopher W. Jenks, *Director, Cooperative Research Programs*
Christopher Hedges, *Manager, National Cooperative Highway Research Program*
B. Ray Derr, *Senior Program Officer*
Charlotte Thomas, *Senior Program Assistant*
Eileen P. Delaney, *Director of Publications*
Ellen M. Chafee, *Editor*

NCHRP PROJECT 03-103 PANEL
Field of Traffic—Area of Operations and Control

Peter Koonce, *City of Portland, Portland, OR* (Chair)
Mark Luszcz, *Delaware DOT, Smyrna, DE*
Woody L. Hood, *Sabra, Wang & Associates, Inc., Columbia, MD*
Susan Langdon, *Savant Group, Dallas, TX*
Bill J. Shao, *City of Los Angeles DOT, Los Angeles, CA*
Aleksandar Stevanovic, *Florida Atlantic University, Boca Raton, FL*
James R. Sturdevant, *Indiana DOT, Indianapolis, IN*
Eddie Curtis, *FHWA Liaison*
Douglas Noble, *Institute of Transportation Engineers Liaison*
Richard A. Cunard, *TRB Liaison*

FOREWORD

By B. Ray Derr
Staff Officer
Transportation Research Board

In addition to covering basic and advanced signal timing concepts, this second edition of the *Signal Timing Manual* addresses establishment of a signal timing program including setting multimodal operational performance measures and outcomes, determining staffing needs, and monitoring and maintaining the system. Some of the advanced concepts addressed include the systems engineering process; adaptive signal control; preferential treatment (e.g., rail, transit, and emergency vehicles); and timing strategies for oversaturated conditions, special events, and inclement weather. The manual will be useful to traffic engineers and signal technicians at any agency operating traffic signals.

In 2008, the FHWA published the *Traffic Signal Timing Manual*, providing a basic synthesis of signal timing practices in the United States. The manual covered fundamental signal timing related to intersection design, vehicle detection, and coordination of signalized intersections; but there were many concepts that, due to resource constraints, were not addressed in great detail. Under NCHRP Project 03-103, Kittelson & Associates, Inc., expanded and upgraded this manual, addressing many of the shortcomings.

Several goals guided the development of the manual: (1) provide material useful to agencies in documenting their own signal timing policies and practices, (2) facilitate the training of new staff, (3) facilitate implementation of more advanced signal timing concepts where appropriate, (4) include examples that illustrate application of the manual material to real-world intersections and systems, and (5) include references that promote a fuller understanding of topics.

During the course of the project, the research team reviewed the pertinent literature (including agency signal timing manuals) and conducted four multi-agency focus groups and a workshop to get feedback on the utility of the 2008 *Traffic Signal Timing Manual* to typical users. From these efforts, the team identified topics that should be expanded or deepened. The user feedback also greatly influenced the organization of the second edition of the manual to increase its usefulness to practitioners.

A PowerPoint presentation providing an overview of this second edition of the *Signal Timing Manual* can be found on the TRB website by searching on *NCHRP Report 812*.

AUTHOR ACKNOWLEDGMENTS

There were numerous contributions, large and small, but all very important to this second edition of the *Signal Timing Manual*. A great TRB project begins with a great panel, and the panel members on this project were both engaged and patient. The project also included focus groups that critiqued the first edition. While the focus groups were complimentary of the contributions made by the first edition, they challenged the project team to tighten up the document (i.e., no repetitive or "nice to know" material), expand the material on advanced concepts (which were only mentioned in the first edition), and update the graphics.

The project team produced an expanded outline, which was used to develop the initial draft chapters. These chapters were reviewed individually by the panel as they were produced. Based on their comments and additional review by the writing team, a second complete draft was provided to the panel and some select outside reviewers. Again, many insightful suggestions were provided and some errors caught. The final product is immensely better due to the thoughtful review by the many contributors below.

FOCUS GROUP PARTICIPANTS

Nader Ayoub, *Iteris, Austin, Texas*
Don Cashdollar, *Florida Department of Transportation*
Larry Colclasure, *Texas Department of Transportation*
John Deskins, *City of Kennewick, Washington*
Hazem El-Assar, *Orange County, Florida*
Chen Hsu, *Virginia Department of Transportation*
Ken Jacobs, *Pinellas County, Florida*
Sage Kamiya, *Sarasota/Manatee County, Florida*
Mike Kinney, *Montgomery County, Maryland*
Rob Klug, *Clark County, Washington*
Aii Mozdbar, *City of Austin, Texas*
Eric Nelson, *Advanced Traffic Solutions, Houston, Texas*
Keith Orr, *City of Portland, Oregon*
Keith Riniker, *Sabra Wang, Baltimore, Maryland*
Tiffany Slaughter, *Oregon Department of Transportation*
Mark Titus, *City of Richardson, Texas*
Nhan Vu, *Virginia Department of Transportation*
Charlie Wetzel, *Seminole County, Florida*
Paul Zebell, *City of Portland, Oregon*

OUTSIDE REVIEWERS

John Black, *City of Richardson, Texas*
Denny Eyler, *SRF Consulting, Plymouth, Minnesota*
Airton Kohls, *University of Tennessee*
Michael Kyte, *University of Idaho*
Paul Olson, *FHWA (retired)*
Wasim Raja, *District of Columbia Department of Transportation*
Robert Saylor, *City of Richardson, Texas*

CONTENTS

CHAPTER 1
INTRODUCTION

CONTENTS

LIST OF EXHIBITS

CHAPTER 1. INTRODUCTION

1.1 FOCUS FOR THE SECOND EDITION

The first edition of the *Traffic Signal Timing Manual* (TSTM) was written as a comprehensive guide for engineers and technicians about signal timing principles, practices, and procedures. The second edition (STM2) is a standalone document that was developed based on TSTM user feedback. Key features of this edition include

- Focused information written for new practitioners and those desiring a better understanding of signal timing fundamentals.
- Addition of four new chapters for more advanced users.
- Material organized so that it is presented once and referenced as needed elsewhere in the document.
- Inclusion of essential information only (i.e., no "nice to know" information).
- References to other documents, instead of repeated material.
- Expanded use of graphics to aid in the explanation of more complex topics.

The second edition of the Signal Timing Manual (STM2) focuses on system users.

The STM2 has an increased focus on signal system users and their priorities. Current signal timing models tend to provide a one-size-fits-all approach to signal timing, which often leads to the incorrect assumption that the model provides the optimum solution. A traffic analyst simply inputs the data required by the model, hits the optimize button, and gives the optimized results to the appropriate person for implementation. The results largely reflect the model's priorities (generally some version of vehicle delay) for system users, which may or may not fit the needs of the actual operating environment or users (including pedestrians, bicycles, and transit).

Performance measures, such as vehicle delay, are often mistaken as the operational objectives.

The STM2 introduces an outcome based approach to signal timing (summarized in Exhibit 1-1), which allows the practitioner to develop signal timing based on the operating environment, users, user priorities by movement, and local operational objectives. Performance measures are then used to assess how well the objectives are being met. Once the objectives and performance measures are established, timing strategies and timing values can be chosen. The final steps of the process involve implementation and observation (i.e., determining if the timing strategies and values are working), as well as sustaining operations that meet the operational objectives through monitoring and maintenance.

*One signal timing objective does **not** fit all conditions.*

This process was developed with an understanding that there is ***not*** a one-size-fits-all method for signal timing. The approach is described in detail in Chapter 3, but brief descriptions of the eight steps in the outcome based process and associated considerations are provided below.

Step 1: Define the Operating Environment

Signal timing should reflect the character of the timing location, so the outcome based approach begins with an assessment of the operating environment. The operating environment goes beyond physical location characteristics and also includes goals of the local operating agency and its regional stakeholders.

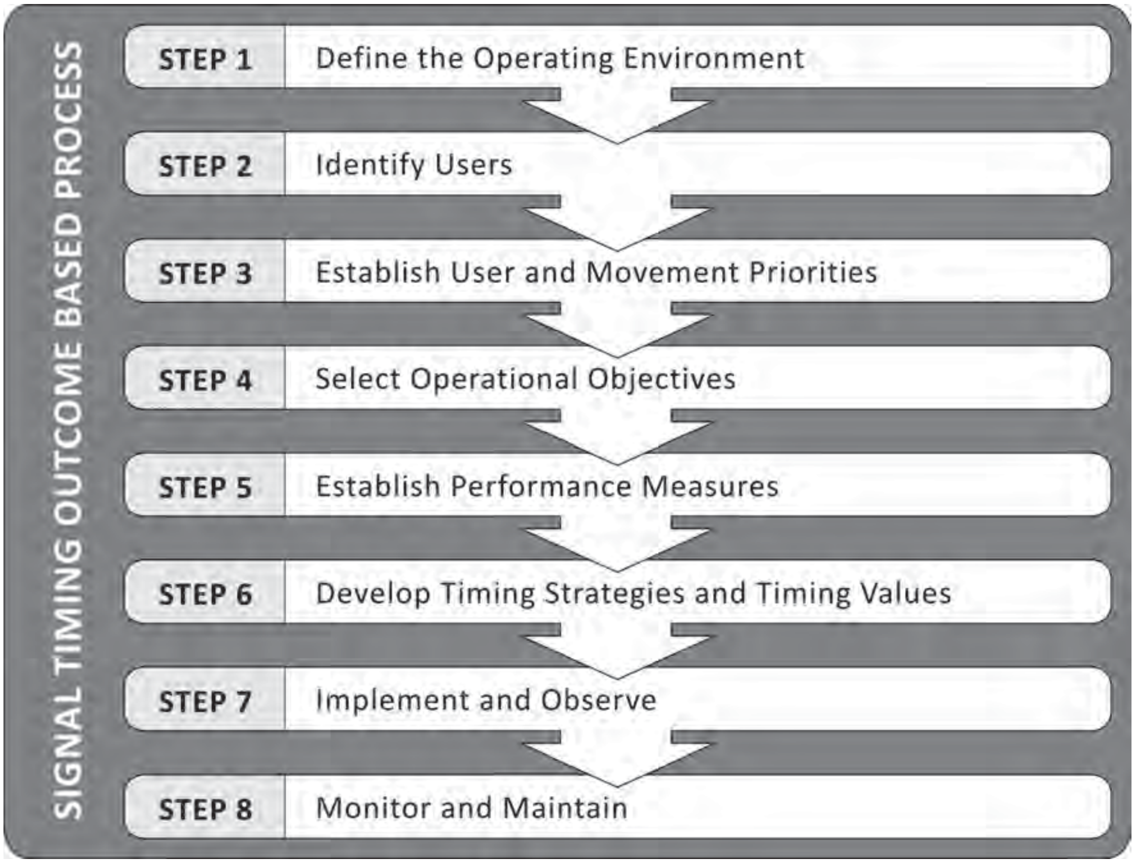


Exhibit 1-1 Signal Timing Outcome Based Process

Step 2: Identify Users

The process continues with the identification of primary users at the focus intersections. This approach allows all users (people on foot, riding on bikes, riding transit, driving trucks, and driving cars) to be considered in the signal timing process.

Step 3: Establish User and Movement Priorities

Priorities should reflect the local operating agency and regional stakeholder goals for mobility. Priorities should be established by movement for the primary users and by location and time of day. For example, in a central business district (CBD), pedestrians might have the highest priority, while in a suburban environment, through vehicle movements on an arterial might have the highest priority during peak hours and a lower priority off-peak.

Step 4: Select Operational Objectives

Once priorities are established, the process requires the establishment of operational objectives (e.g., pedestrian safety or vehicle mobility) by location and time of day. Non-vehicle-oriented operational objectives are often more difficult to assess because of their qualitative nature; however, performance measures can be selected for qualitative as well as quantitative assessment.

Step 5: Establish Performance Measures

Traditional optimization tools generally focus only on simple vehicle-oriented performance measures because they are easy to quantify and, therefore, easy to "optimize." However, vehicle stops and delay may be less important than transit and pedestrians in a CBD, as well as other existing or developing areas with significant pedestrian, bicycle, and transit activity. The practitioner needs to make appropriate adjustments to the traffic signal timing process to account for the operating environment and user priorities.

Step 6: Develop Timing Strategies and Timing Values

Once operational objectives and their associated performance measures have been determined, the process continues with the development of signal timing strategies (i.e., minimizing cycle length or favoring arterial through traffic) and selection of appropriate timing values.

Step 7: Implement and Observe

The next step is implementing the signal timing values and making final adjustments to the timing parameters.

Step 8: Monitor and Maintain

After implementation, a successful program requires ongoing monitoring and maintenance. Collecting periodic (at least annual) volume data at a mid-block location on each arterial or subsystem is essential when determining if shifts in traffic characteristics occurred or if further investigation is necessary. A good maintenance management system can help an agency identify issues beyond the signal controller, including communication and detection issues, which are often significant contributors to poor operations.

1.2 STM2 ORGANIZATION

The outcome based approach not only guides the development and implementation of signal timing plans, but also guides the organization of the STM2. As shown in Exhibit 1-2, the manual has 12 chapters.

Chapters 1 through 4 address signal timing fundamentals, providing guidance on the elements used to support a successful signal timing program, as well as concepts necessary for understanding modern traffic signal operations. Because a poorly designed signal can never operate at its full potential, Chapter 4 focuses specifically on signal design.

Chapters 5 through 8 are focused on basic signal systems, presenting the detailed, practical information necessary to appropriately time a traffic signal controller and to operate it in a system with multiple controllers. These chapters include important information about controller features that are often overlooked because they are not included in signal timing optimization tools. The information presented in Chapters 5 through 8 is intended to help new practitioners achieve locally appropriate results and to serve as a reference for more experienced practitioners.

Guidance on advanced signal systems and applications is provided in Chapters 9 through 12. The material begins with a brief introduction to Systems Engineering, a process that reduces risk by clearly defining needs and the specific requirements to meet local needs. These chapters conclude with specific information about managing competing priorities in a multimodal system.

Exhibit 1-2 STM2 Organization

Part	Chapter	Title
SIGNAL TIMING FUNDAMENTALS	CHAPTER 1	Introduction
	CHAPTER 2	Signal Timing Program
	CHAPTER 3	Signal Timing Concepts
	CHAPTER 4	Signal Design
BASIC SIGNAL SYSTEMS	CHAPTER 5	Introduction to Timing Plans
	CHAPTER 6	Intersection/Uncoordinated Timing
	CHAPTER 7	System/Coordinated Timing
	CHAPTER 8	Implementation and Maintenance
ADVANCED SYSTEMS AND APPLICATIONS	CHAPTER 9	Advanced Signal Systems
	CHAPTER 10	Preferential Treatment
	CHAPTER 11	Special Conditions
	CHAPTER 12	Oversaturated Conditions

It should be noted that the STM2 relies on a number of important reference documents for details. It is intended to complement policy documents such as the *Manual on Uniform Traffic Control Devices for Streets and Highways* (MUTCD, *1*) and is not intended to replicate or replace the *Highway Capacity Manual* (*2*), or national or local engineering documents on signal timing, nor is it intended to serve as a standard or policy document. The manual has been structured and written as concisely as possible, and because the manual only references these external documents, it should not become out-of-date as those documents are updated. Practitioners are cautioned to consult current documents for changes. For example, documents such as the MUTCD (*1*) are periodically updated, and some guidance in this document is based on the 2009 edition.

1.3 REFERENCES

1. *Manual on Uniform Traffic Control Devices for Streets and Highways,* 2009 Edition. United States Department of Transportation, Federal Highway Administration, Washington, D.C., 2009.

2. *Highway Capacity Manual 2010.* Transportation Research Board of the National Academies, Washington, D.C., 2010.

CHAPTER 2
SIGNAL TIMING PROGRAM

CONTENTS

LIST OF EXHIBITS

CHAPTER 2. SIGNAL TIMING PROGRAM

*While effective signal timing is necessary, it will not **automatically** sustain a successful signal timing program.*

Signal timing is the process of selecting appropriate values for timing parameters implemented in traffic signal controllers and associated system software. Effective signal timing programs ensure that signal timing parameters are appropriate over the ***life*** of the traffic signal system, by monitoring all aspects of traffic signal implementation, operations, and maintenance consistent with community needs. A successful program requires agency staffing and maintenance funding that is consistent with the level of service planned.

2.1 ELEMENTS OF SUCCESSFUL SIGNAL TIMING PROGRAMS

Each operating agency has common traits (described in detail throughout this section) that can increase the likelihood of a signal timing program receiving support from decision-makers. In general, effective signal timing programs tend to have:

- ***Effective intra-agency and inter-agency cooperation***, which fosters knowledge-sharing, access to resources, and a higher level of customer service.
- ***Internal champions and support*** from leadership within the program.
- ***External support*** from elected leaders and stakeholders.
- ***Agency goals and desired outcomes*** for the signal system.
- ***Structured programs*** for tasks such as signal timing, performance measurement, maintenance, training, and outreach.

2.1.1 Leadership

It has been well documented that leadership is the most important factor for a successful program. Leadership often starts with a champion at one or more organizations and, possibly, at one or more levels of an organization. There are many different leadership approaches that can result in a successful program, especially given different organizational structures.

2.1.2 Self-Assessment and Evaluation

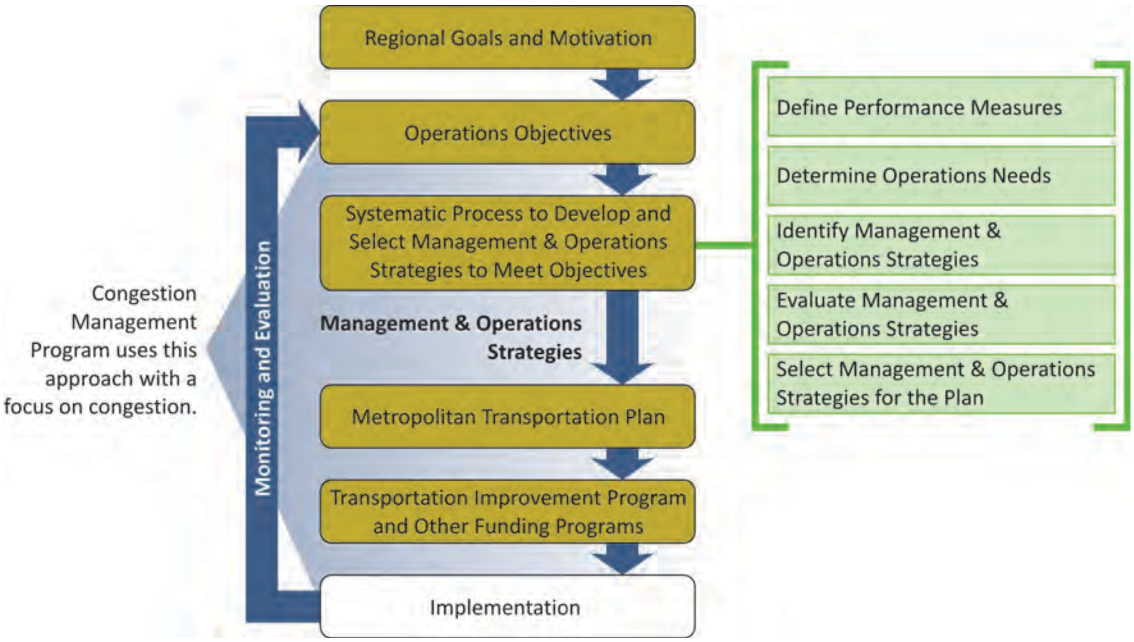
Self-assessment can help agencies understand what works, what could be improved, and what keeps the agency operating. If this is done collaboratively throughout a region, it will be easier to develop a shared vision. Resources for self-assessment include (but are not limited to) the *Traffic Signal Self-Assessment* (www.ite.org/selfassessment/TSOSelfAssessment11.pdf), the *Traffic Signal Audit Guide* (*1*), Federal Highway Administration (FHWA) peer review assistance, and a multitude of other FHWA documents that support successful signal timing programs.

2.1.3 Funding Mechanisms

Funding is an essential part of a signal timing program and is often available through a variety of sources such as direct agency funding, state-local arrangements, public-private partnerships, and federal funding. In order to be successful, a program must acquire enough funding to serve its system users (*2*). Identifying and documenting agency needs for funding organizations, like a Metropolitan Planning Organization (MPO), or local policymakers is part of sustaining a traffic signal timing program.

One funding opportunity is increasing the focus on management and operations (M&O) in a Metropolitan Transportation Plan (MTP) (*3*). For example, an MTP can include a congestion management process (like that shown in Exhibit 2-1), which is a potential source of funding. This objectives-driven approach is consistent with the outcome based process introduced in Chapter 1.

Exhibit 2-1 Congestion Management Process



Source: Adapted from *Advancing Metropolitan Planning for Operations: An Objectives-Driven, Performance-Based Approach: A Guidebook* (*3*).

2.1.4 Training Programs

By fostering a true understanding of a program, individuals are empowered to find solutions and to sustain the technical aspects of a program.

Training is an ongoing need because staff (at all levels) need to understand the goals and objectives of a signal timing program, as well as acquire the skills necessary to accomplish their assignments. Training of various kinds is available from equipment vendors, software providers, universities, states, the United States Department of Transportation, and the National Highway Institute. *Traffic Signal Operations and Maintenance Staffing Guidelines* (*4*) provides a comprehensive overview of staffing for signal operations, with additional material available in Chapter 8.

2.1.5 Public Involvement and Outreach

Good communication is necessary in acquiring public and political support for a sustainable signal timing program. The means of communication should be appropriate for the agency. It could be as simple as providing a phone number on the side of a traffic signal cabinet, or it could be as detailed as posting an explanation of how signals are being timed on the agency website. As performance measures are collected, they can also be used to indicate progress or gather support for improvements.

2.2 BENEFITS OF REGIONAL SIGNAL TIMING PROGRAMS

By bringing a diverse set of strengths together, a regional signal timing program can provide added value to roadway users. Regional programs produce more efficient and consistent operations than individual programs working alone; regional programs can lead to improved mobility and safety across a region. Specifically, some benefits of a regional program may include (*5*)

- ***Advancement of projects*** that are too large for a single agency to undertake but are manageable as regional or state transportation improvements.
- ***Increased access to funding*** through joint applications.
- ***Availability of a central point of contact***, which simplifies information processing and sharing of feedback from stakeholders and users.
- ***Leveraging resources and experience*** through shared training, office space, equipment purchases, technician support, and Information Technology (IT)/Information Systems (IS) staff, which often results in a more collaborative working environment.
- ***Consistent signal operations***, resulting from practitioners having the same training or certification and using the same guidance on signal timing parameters.
- ***Improved signal operations***, resulting from better cross-agency coordination, timing practices, and shared knowledge.
- ***Smoother traffic management during special conditions***, through shared resources and better communication when shifting traffic from one agency's facilities to another.

More information can be found in *The Collaborative Advantage: Realizing the Tangible Benefits of Regional Transportation Operations Collaboration: A Reference Manual* (*6*), *Scan 07-04: Best Practices in Regional, Multiagency Traffic Signal Operations Management* (*7*), and *NCHRP Synthesis 420* (*8*).

2.3 REFERENCES

1. *Traffic Signal Audit Guide*. National Transportation Operations Coalition, Washington, D.C., 2007.
2. Denny, Jr., R. W. *Improving Traffic Signal Management and Operations: A Basic Service Model*. Report FHWA-HOP-09-055. Federal Highway Administration, United States Department of Transportation, 2009.
3. Grant, M., J. Bauer, T. Plaskon, and J. Mason. *Advancing Metropolitan Planning for Operations: An Objectives-Driven, Performance-Based Approach: A Guidebook*. Report FHWA-HOP-10-026. Federal Highway Administration, United States Department of Transportation, 2010.
4. Gordon, R., and C. Braud. *Traffic Signal Operations and Maintenance Staffing Guidelines*. Report FHWA-HOP-09-006. Federal Highway Administration, United States Department of Transportation, 2009.

Signal Timing Manual, Second Edition

5. Koonce, P., K. Lee, and T. Urbanik. *Regional Traffic Signal Operations Programs: An Overview*. Report FHWA-HOP-09-007. Federal Highway Administration, United States Department of Transportation, 2009.
6. Bauer, J., M. Smith, and A. Armstrong. *The Collaborative Advantage: Realizing the Tangible Benefits of Regional Transportation Operations Collaboration: A Reference Manual.* Report FHWA-HOP-08-001. Federal Highway Administration, United States Department of Transportation, 2007.
7. Jennings, B., G. Bachmann, E. Curtis, S. Misgen, V. Nguyen, J. B. Renick, and K. N. Balke. *Scan 07-04: Best Practices in Regional, Multiagency Traffic Signal Operations Management*. Final Report, NCHRP Project 20-68A. American Association of State Highway and Transportation Officials, 2013.
8. Balke, K. N., and A. Voight. *NCHRP Synthesis of Highway Practice 420: Operational and Institutional Agreements That Facilitate Regional Traffic Signal Operations*. Transportation Research Board of the National Academies, Washington, D.C., 2011.

CHAPTER 3
SIGNAL TIMING CONCEPTS

CONTENTS

Signal Timing Manual, Second Edition

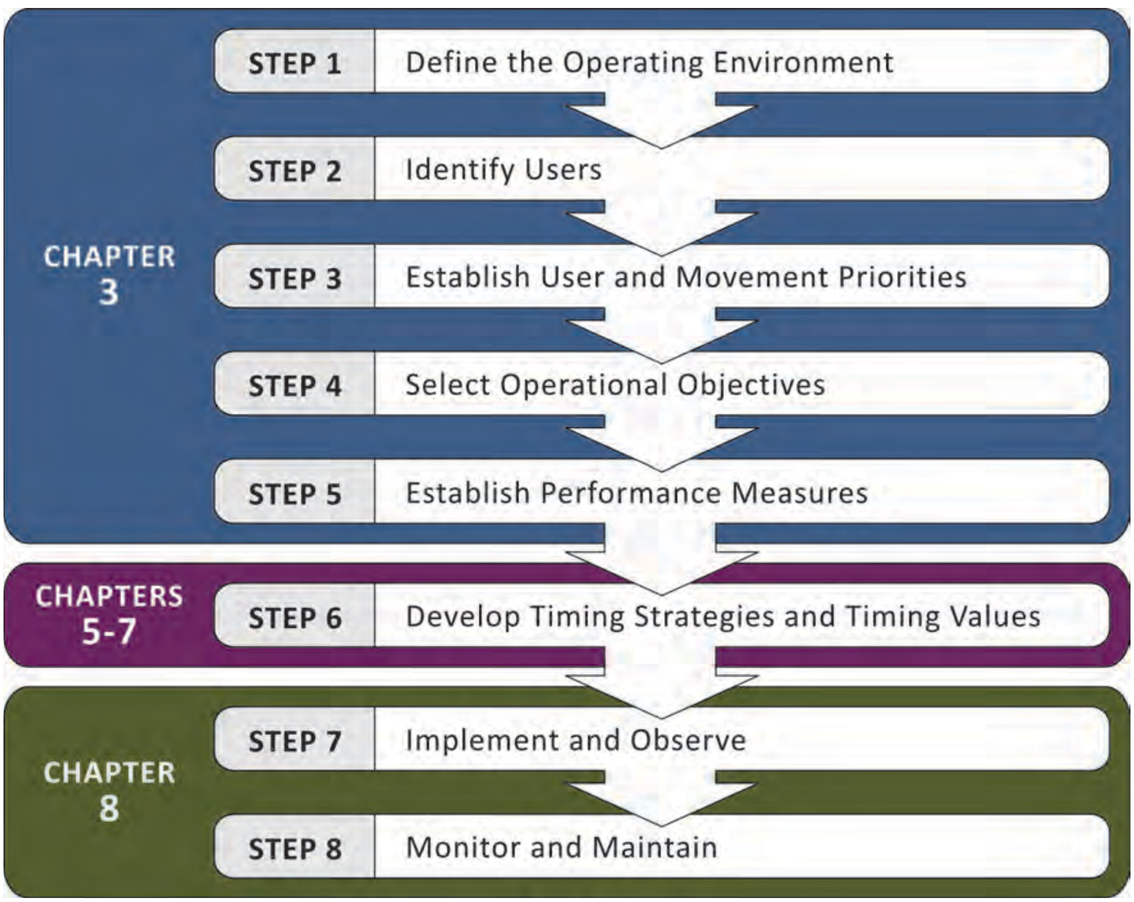
LIST OF EXHIBITS

CHAPTER 3. SIGNAL TIMING CONCEPTS

Operating environment, users, and movements are the building blocks for developing a signal timing plan.

This chapter provides an overview of signal timing basics, organized using the outcome based process introduced in Chapter 1 (and shown again in Exhibit 3-1). The outcome based process is a modern approach to signal timing and reflects the complex nature of many signalized intersections and operating environments. It encourages practitioners to consider all system users (e.g., pedestrians, bicycles, motor vehicles, emergency vehicles, transit, and rail) and to establish user priorities by movement for each signalized intersection location. The focus is applying those elements to the selection of operational objectives that clearly define desired signal timing outcomes. In order to assess the effectiveness of a signal timing plan, performance measures must also be identified for each objective. Performance measures can be used to determine initial success, as well as to monitor and sustain desired outcomes throughout the life of a signal. Traffic signal timing adjustments should be part of a continual process of adapting to changing conditions, not the result of retiming once in a long while.

Exhibit 3-1 Chapter Outline Using the Outcome Based Process



Before discussing the outcome based process, this chapter introduces the basic components of a traffic signal system. Modern traffic signal controllers are essentially computers that accept requests for service and control user displays based on the rules established by the practitioner. The specific timing parameters have a direct relationship to service priority. Therefore, it is necessary to understand the basic operation of a modern traffic signal controller before assessing various alternative timing strategies and values.

The second part of this chapter explains the signal timing outcome based process. As shown in Exhibit 3-1, detailed information about the eight steps is divided among several chapters. The first five steps essentially define desired outcomes and are discussed in this chapter. Chapters 5–7 provide the details on how to develop timing strategies, intersection timing parameters, and system timing parameters. Chapter 8 discusses the critical steps of implementation, initial observation, long-term monitoring, and maintenance. The last step is part of continual improvement, which feeds back to earlier steps as needed. Not only can operational objectives be maintained through such efforts, but the necessity of a signal timing program (discussed in Chapter 2) may become more apparent. For example, if a previous reduction in maintenance resulted in increased stops, delay, and fuel consumption, then a stronger case may be made for increased maintenance moving forward.

This chapter explains the steps leading up to the selection of signal timing values. There are many initial pieces of information that must be considered if a signal timing plan is to be effective.

3.1 TRAFFIC SIGNAL BASICS

A few terms warrant introduction because of the potential for multiple interpretations. Terms are generally defined when first used, but a glossary is available at the end of the manual.

- ***User:*** A specific category of persons receiving service at a traffic signal. Users include pedestrians, bicycles, passenger cars, trucks, emergency vehicles, and transit. Users (which in aggregate constitute the “traffic” at a signalized intersection) are the primary focus of the outcome based process.
- ***Vehicle:*** A general term that encompasses all devices, without regard for the user, including passenger cars, bicycles, trucks, and buses. Traditionally, vehicles have been the primary unit of measure for most traffic signal analysis (e.g., vehicle volume or vehicle delay). Vehicles are sometimes categorized as light (passenger cars and light trucks) and heavy (heavy trucks and buses), and in order to account for their operational differences, vehicle equivalents are commonly applied (i.e., passenger car equivalents for heavy vehicles). In this manual, the term vehicle is used when it is necessary to include multiple user types in the description of traffic signal operations (e.g., a green indication serving passenger cars, trucks, bicycles, and buses) without regard for their specific needs.
- ***Movement:*** A term that describes user actions at an intersection (e.g., northbound vehicular left turn or pedestrian using the west crosswalk). Movements can be ***permitted***, requiring users to yield to others when given a green indication, or ***protected***, which gives users the right-of-way without any conflicts.
- ***Practitioner:*** The person determining signal timing.

3.1.1 Common Signal Components and Interactions

Each signalized intersection has common components that provide the basis for signal operations (illustrated in Exhibit 3-2): a controller, a cabinet, displays (or indications), and, typically, detection. Exhibit 3-3 depicts the basic interaction among these signalized elements and the system users. It is important for a practitioner to recognize the relationship among users, detection, and the signal controller, as well as how that relationship influences the resultant displays (i.e., vehicle displays, pedestrian displays, and occasionally special displays for specific users) to ultimately inform users of their right-of-way. Only a basic understanding of the components is required for this chapter. However, it should be noted that in Exhibit 3-2, the vehicle displays (typically for passenger cars, trucks, and buses) also control the bicycles, and the intersection may or may not have bicycle detection and timing to control the shared vehicle displays.

Exhibit 3-2 Common Signalized Intersection Components

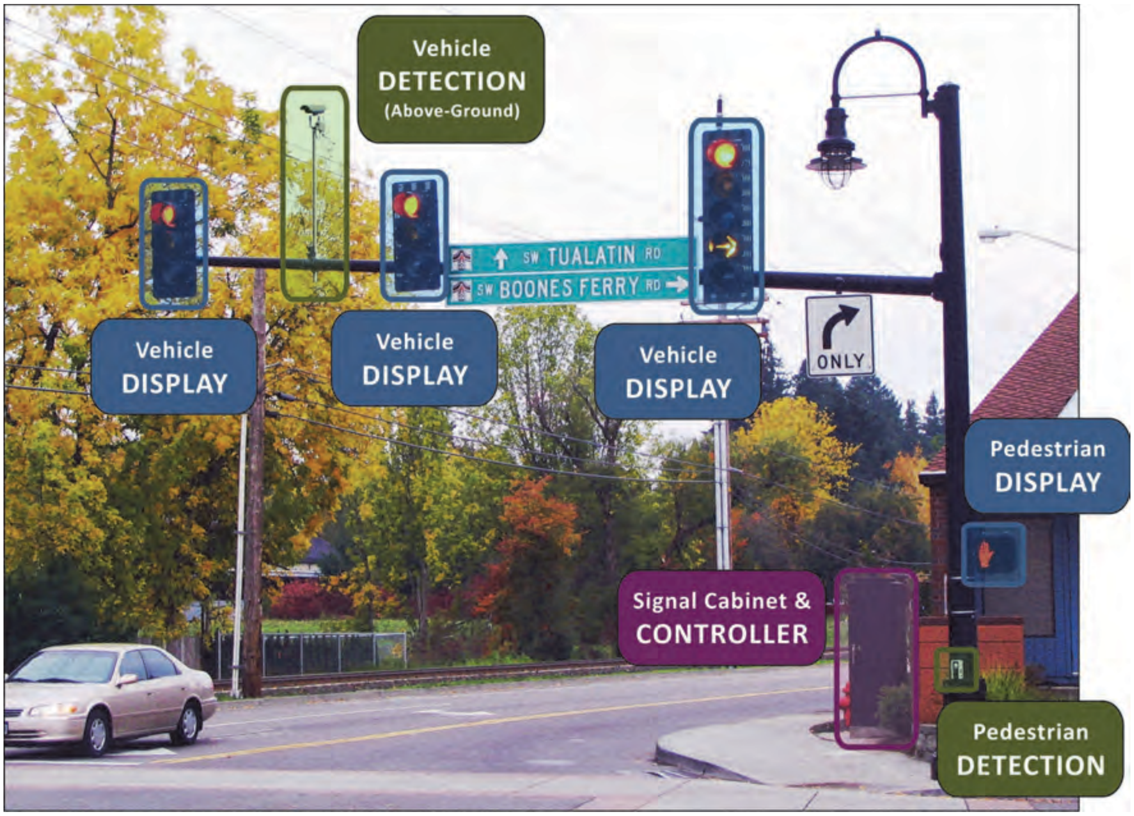
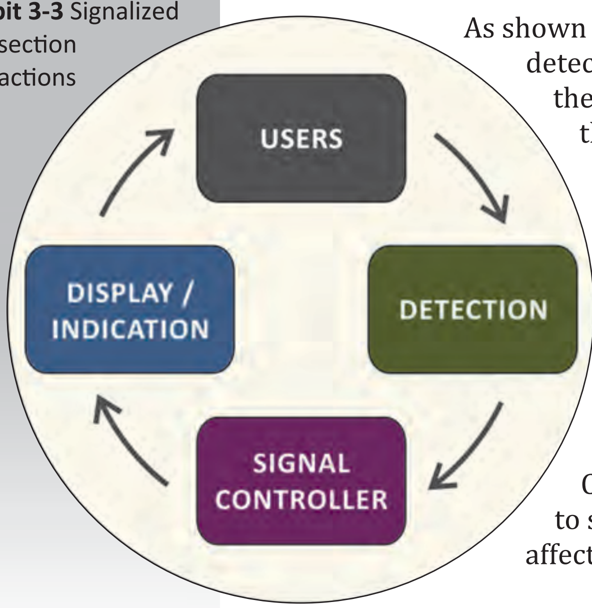


Exhibit 3-3 Signalized Intersection Interactions



As shown in Exhibit 3-3, the user is part of a continuous process. The detectors (e.g., vehicle, pedestrian, or bicycle) send messages to the controller as users approach the intersection. The controller then uses the detector inputs to change the user displays (typically vehicle and pedestrian) based on the signal timing parameters defined by the practitioner. If many users approach the intersection at once, the process becomes more complicated, and the controller must make decisions (based on the timing parameters that control user priorities) about which movements will receive the right-of-way. The way that the controller assigns time to users is highly dependent on the detection and signal timing parameters (controller settings) that have been programmed. Chapter 4 discusses the critical design considerations related to signal hardware (i.e., cabinets, controllers, and detection) that affect the type of operation that is feasible.

3.1.2 Basic Signal Controller Concepts

Signal controllers (specifically the internal program called firmware) are an essential component of traffic signal control. While the remainder of this manual summarizes various functionalities, parameters, and values that can be programmed into signal controllers, this section only introduces basic controller concepts. There are two defining characteristics of signal controllers: (1) how the signal controller interprets demand at the local intersection and (2) how the signal controller relates to other controllers.

3.1.2.1 Interpretation of Local Demand

Signal controllers can adjust operations using external information about user demand (i.e., requests for service). It is the type of external information that the controller utilizes that defines the type of signal control (as outlined in Exhibit 3-4). Fully-actuated signal control is fully adaptive to local traffic conditions because it utilizes demand information from detectors (at least vehicle and pedestrian) located on all approaches. Fully-actuated intersections offer the most flexibility. Traffic signal controllers that utilize detection only in some lanes are considered semi-actuated because they can only adapt operations based on partial demand information.

A controller that does not use any detection to adapt operations is considered pretimed. Modern pretimed traffic signal systems are used in some special cases. Pretimed systems are the least costly to build because they do not have the expense associated with detection. They use signal timing values that were calculated and programmed into the controller based on historical data. These types of systems are typically used in central business districts (CBDs) with closely spaced intersections (often in grid networks), where detection has little potential to improve operations given the need to maintain well-defined relationships between intersections.

Exhibit 3-4 Types of Signal Control

3.1.2.2 Relationship to Other Controllers

Regardless of the type of detection being used, signal controllers can operate either alone or as part of a system. Uncoordinated signal timing allows the intersection to run independently (or “free”) of any other intersection, while coordinated timing operates several signals as a system. It is often believed, incorrectly, that detection cannot be used to influence the coordinated phases when an intersection is coordinated. Modern controllers can actuate a portion of the coordinated phases, providing additional flexibility. Detailed information about coordinated timing is available in Chapter 7.

Even if operated as part of a system, the local intersection is still an active part of operations. Traditional traffic signal systems (that use coordination to minimize stops for through traffic) impose a common cycle length to maintain a consistent relationship between adjacent traffic signals. Coordination essentially constrains the local traffic signal controller to a timing plan that will achieve the operational objective of corridor progression (resulting in fewer stops along the corridor).

Adaptive systems are an alternative to traditional coordination and use different algorithms to adjust signal timing parameters. Adaptive systems still rely on local controllers for many timing parameters and share most features of traditional systems, but they have more flexibility in how they adjust timing parameters. Adaptive systems that have advanced system capabilities are discussed in Chapter 9. Adaptive systems should only be considered after the capabilities of traditional systems have been fully considered. Chapters 5–7 focus on traditional traffic signal control, which is currently used for the vast majority of systems and is more than adequate for most locations.

3.2 INITIAL SIGNAL TIMING CONSIDERATIONS

This section will give the practitioner a general idea of the type of information that should be gathered for making informed signal timing decisions. Exhibit 3-5 summarizes the basic information that a practitioner might investigate as part of the first three steps of the outcome based process. Detailed information about each of these topics is provided in the following sections.

Exhibit 3-5 Initial Signal Timing Considerations

Outcome Based Process		Initial Signal Timing Considerations
STEP 1 Define the Operating Environment	Multi-Jurisdictional Impacts	□ Is the system of signals located in a single or multiple jurisdictions? □ If multiple, is signal timing performance consistent across jurisdictional boundaries? □ Is there an existing agreement in place that defines certain signal timing parameters? □ Would an agreement between jurisdictions to establish consistent signal timing strategies be beneficial?
	Location and Associated Environment	□ Is the intersection or system of intersections located in a rural, suburban, or urban area? □ Is the area transitional or undergoing changing land uses?
	Roadway Classification	□ How are the roadways classified (major street and minor street) where the signal system is located (e.g., freeway interchange, major arterial, minor arterial, major collector, minor collector, or local street)? □ Are there specific pedestrian, bicycle, freight, or transit route needs? □ How does the classification affect user expectations of the facility?
	Transportation Network	□ How closely spaced are the signalized intersections? □ Is there a reason to consider multiple intersections as a system when developing the signal timing? □ Is the minor street coordinated? □ Is there nearby rail (freight or passenger) requiring preemption?
STEP 2 Identify Users		□ What is the existing mix of users (e.g., pedestrians, bicycles, light vehicles, heavy vehicles including trucks and transit vehicles, priority vehicles, and rail including freight and passenger)? □ Does the mix change by time of day? □ Are there unique travel patterns?
STEP 3 Establish User and Movement Priorities		□ Who are the critical users at the intersection(s)? □ Does the critical user change by time of day? □ Does the jurisdiction have any policies related to user priorities? □ What are the critical movements?

3.2.1 Define the Operating Environment

The first step in the outcome based process is defining the operating environment. Signals timed together as a system should have a common operating environment so that the practitioner can work from consistent priorities and objectives when developing signal timing plans.

Conducting a site visit and/or gathering information (i.e., local knowledge) is important for correctly characterizing the operating environment and validating desired outcomes.

3.2.1.1 Multi-Jurisdictional Impacts

In some instances, a system of signals may operate across multiple jurisdictions (e.g., a system located along a corridor that runs from within a city's jurisdiction to another city's, county's, or state's jurisdiction). If a system is close to or across a jurisdictional boundary, the operating agency should coordinate with all affected jurisdictions to discuss and develop an agreement on objectives and various signal timing parameters, as appropriate.

System users will not expect signal timing performance to vary across jurisdictional boundaries, so every effort should be made to time the signals in a way that allows a seamless transition. Maintaining consistent timing values (e.g., cycle length and clearance times) across jurisdictional boundaries will help users know what to expect. For example, a consistent cycle length will allow cross-jurisdictional coordination, thereby reducing stops. The practitioner should verify which signal timing values he or she can change and which should only be adjusted within the jurisdiction agreement.

Signal timing plans should be transparent to users across jurisdictional boundaries.

3.2.1.2 Location and Associated Environment

One of the primary factors affecting signal timing is the environment at the intersection (or system of intersections). In particular, the practitioner will want to identify whether the intersections are located in a rural, suburban, or urban environment because each location will require the practitioner to consider different priorities for timing objectives.

The objectives for signal timing will be different depending on if a signal system is located in a rural, suburban, or urban environment. Objectives may also change by time of day or by traffic volume.

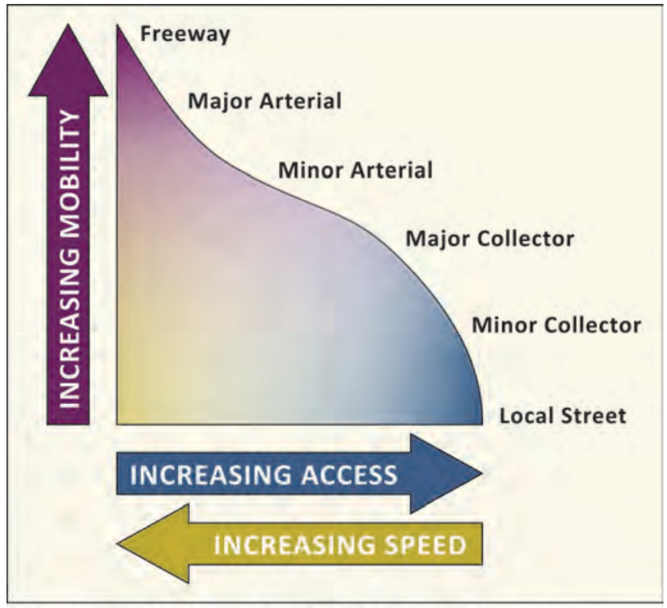
Signal timing objectives for intersections located in rural environments generally focus on the existence and use of setback detection, as well as minimizing the number of drivers experiencing decision zones. Decision zones (and associated detection) are described in detail in Chapter 4, but are locations where different drivers may make different decisions, resulting in potential conflicts.

Signal timing in suburban areas often focuses on achieving smooth flow (minimizing stops) along arterials. This can be achieved through strategies that include coordinating intersections and actuating the pedestrian movements (if pedestrian crossing demand is low during certain or all times of day). Timing objectives may change by time of day to reflect changing traffic patterns, and having good detection allows the most flexibility.

The primary focus for signal timing in downtown urban environments is typically accommodating all user groups (e.g., pedestrians, bicycles, passenger cars, and transit). In order to manage vehicle queues, a shorter cycle length may be chosen that reflects the shorter distance between intersections. Other multimodal strategies include setting the progression speed based on bicycle travel speeds, setting the recall mode based on pedestrian priority, and defining offsets to prioritize transit vehicle operations.

3.2.1.3 Roadway Classification

Exhibit 3-6 Roadway Classifications



Source: Adapted from *A Policy on Geometric Design of Highways and Streets* (*1*)

The roadway facility within the larger transportation system can influence user expectations, which in turn can influence how the roadway should be timed. There is the traditional tradeoff between mobility and access (depicted in Exhibit 3-6), but other factors can impact signal timing settings as well, including

- Freight route designations,
- Key pedestrian crossings,
- Signalized multi-use trail crossings,
- Bike boulevards, and
- Rapid or high-capacity transit routes.

A practitioner should understand the roadway classifications of the transportation network and, equally important, user expectations.

3.2.1.4 Transportation Network Characteristics

The configuration of the transportation network can have a significant impact on the way its traffic signals are timed. The practitioner should consider the effect that nearby intersections will have on one another based on their spatial relationship as well as the desired speed of traffic.

Exhibit 3-7 Individual Intersection

Individual intersections are sometimes referred to as "isolated," which refers to a mode of operation, rather than a spatial relationship. Individual intersections may also be described as "free" operation, which indicates they are not currently being coordinated.

Individual signalized intersection operations (depicted in Exhibit 3-7 and covered in detail in Chapter 6) typically occur with long signalized intersection spacing (a half mile or more) or when other factors (such as late-night operation with low traffic volumes) make independent operations preferable. Individual signalized intersections can be timed without the explicit consideration of other traffic signals, allowing the flexibility to set signal timing parameters that are optimal for the individual intersection. This type of operation is often called "free" (or "isolated") because it operates without need for or benefit from coordination. In these cases, good detection on all approaches is necessary for high operational and safety performance. It should be noted that a signalized intersection may operate both in a free and a coordinated manner depending on traffic conditions. It is not necessary for a signalized intersection to operate in the same manner at all times.

Exhibit 3-8 Arterial Intersections

For signalized intersections located along arterial streets with high volumes of through traffic (depicted in Exhibit 3-8), individual signalized intersection operations can usually be improved by considering coordination of the major street movements. For most arterial streets with signals that are spaced a half mile apart or less, coordinated operations can yield benefits by improving progression between signals. On arterials with higher speeds and/or limited access, it can be beneficial to coordinate signals spaced a mile apart or even further. Additional guidance on coordination is available in Chapter 7.

Exhibit 3-9 Grid Network Intersections

Signalized intersections can also be located in grid networks (depicted in Exhibit 3-9). Grid networks are more difficult to time because tradeoffs have to be made in order to coordinate the intersections. In these cases, the entire network is often timed together to ensure consistent behavior among intersections. Grid networks, particularly in downtown environments with short block spacing, are frequently timed using pretimed plans (i.e., all phases on vehicle and pedestrian recall) because the same amount of time is generally desired each cycle to maintain consistent relationships between the closely spaced intersections.

Signals that are located very close together (less than 500 feet or 7 to 10 seconds apart) often require signal timing settings and strategies that manage queues (typically by keeping cycle lengths low); however, pedestrians may be a driving force in the timing. In some small, closed networks with closely spaced intersections (e.g., grid of four closely spaced, one-way streets, triangle created by three intersections, or closely spaced offset-T intersections), it may be beneficial to operate intersections with a single controller to tightly control the relationship among all movements.

Geometric constraints associated with any type of transportation network (i.e., a lack of turn lanes, short turn-lane storage bays, or driveway locations) can have important impacts on operations. Signal timing strategies should reflect these location-specific considerations. Potential strategies are discussed in later chapters.

3.2.2 Identify Users

The mix of users and traffic patterns at an intersection will influence operations.

User characteristics clearly influence the effectiveness of signal timing. In particular, the mix of users and their traffic patterns will affect how an intersection operates (as depicted in the multimodal environment in Exhibit 3-10). The following users should be considered during the timing process:

- ***Pedestrians.*** Pedestrians with slower walking speeds (e.g., children and elderly), persons with mobility limitations, and pedestrians with visual impairments need more time to cross the street. Pedestrian walk times and clearance intervals (discussed in Chapter 6) may need to be adjusted to reflect local conditions and/or local policies. A leading pedestrian interval may also be warranted when high pedestrian volumes are competing with a high number of turning vehicles. In all cases, timing should, at a minimum, meet requirements in the *Manual on Uniform Traffic Control Devices* (MUTCD, *2*).
- ***Bicycles.*** High bicycle use at an intersection may warrant special bicycle detection and associated bicycle minimum green times or extension times. Timing parameters related to phase initiation and phase termination (i.e., minimum green, yellow change, and red clearance) are critical to bicycles (see Chapter 6 for more details) and may need to be adjusted to account for their lower speeds and acceleration characteristics.
- ***Light Vehicles.*** The number of light vehicles (i.e., passenger cars and light trucks) using an intersection will impact many signal timing parameters (discussed throughout Chapters 6 and 7), including cycle length, the time allocated to each phase, and the order of the phases.
- ***Heavy Vehicles.*** Truck traffic requires accounting for slower acceleration and longer deceleration times and the larger size of vehicles, which can influence queue storage and detection setting needs. Truck priority (discussed in Chapter 10) should be considered at high-speed or downhill approaches to intersections with high volumes of truck traffic.
- ***Emergency Vehicles.*** Emergency vehicles may justify preferential treatment through the use of preemption and/or priority (discussed in Chapter 10).
- ***Transit Vehicles.*** Transit vehicles may justify special phasing and preferential treatment through the use of preemption and/or priority (see Chapter 10).

Exhibit 3-10
Multimodal Environment

3.2.3 Establish User and Movement Priorities

User and movement priorities should be established prior to development of signal timing plans.

Operating environment and intersection users play a key role in how an intersection (or system of intersections) functions. The practitioner should identify user priorities for the signal timing plan(s), in conjunction with any relevant jurisdictional standards and policies. Priorities may be relative (e.g., pedestrians over vehicles) or absolute (e.g., railroad preemption). For example, if pedestrian volumes are high, pedestrians may be prioritized over vehicles during certain times of the day. In order to establish pedestrian priority during times with high right-turning vehicle traffic, the signal timing plan could provide a leading pedestrian interval (i.e., pedestrians get a walk indication before vehicles get a green indication).

In addition to user priorities, the practitioner should also consider movement priorities. Critical movements can be determined using very simple analyses (e.g., critical movement analysis) or through the use of a variety of traffic operations software. It is important for the practitioner to understand that a software package may not automatically generate the desired solution. The practitioner may need to make adjustments to match the desired outcomes. By observing the critical movements at each intersection, the practitioner can understand how the intersection is currently operating and any potential changes that should be included in the timing plan development (e.g., more time allocated to the minor street to reduce phase failures).

User and movement priorities will influence many signal timing parameters, including the cycle length, time allocated to each phase, and phase order.

Giving priority to any one user group or movement requires tradeoffs for other users and movements. If the priority is arterial vehicular through movements, the priority for smooth flow (or minimization of arterial stops) for those vehicles might mean more delay for minor street vehicles and pedestrians crossing the arterial. Therefore, before selecting user and movement priorities, it is important to collect data at the intersection(s) and identify patterns, including

- Proportion of turning vehicles,
- Corridor directionality,
- Peaking characteristics,
- Lane usage,
- Weekday and weekend characteristics,
- Typical and atypical trends (e.g., incident, crash, construction, game, concert, shift change, or convention), and
- Origin-destinations.

3.3 DATA COLLECTION

An effective assessment of the operating environment, users, and priorities includes data collection at the intersection(s). Required data can typically be separated into the following categories: traffic characteristics, intersection geometry, traffic control devices, existing signal timing, and crash history. If an intersection does not yet exist, then estimates will need to be developed, but, if possible, the practitioner should conduct site visits at the existing intersections during multiple traffic demand scenarios.

The following sections provide detailed explanations of the data that can be collected to aid in signal timing development, specifically for the initial steps in the outcome based process. Exhibit 3-11 summarizes the types of information that may be helpful to a practitioner when determining objectives and performance measures. In many instances, the data described herein may have been collected previously as part of a signal warrant analysis or traffic operations study. Care should be taken when reusing older data to understand and account for its potential inaccuracy. Some traffic data should be regularly collected at the same location along a corridor (typically at a mid-block location free of peak-period queuing) in order to track changes in traffic demand.

Exhibit 3-11 Data Collection

Data Collection Category		Potential Data to Collect	
Field Data		☐ Photos and/or video of the intersection(s) ☐ Observations from adopting the role of different types of users ☐ Upstream and downstream bottleneck considerations	☐ Observations about operations watched from a stationary position ☐ Potential impacts from timing changes
Traffic Counts	24-Hour Weekly Counts	☐ Tube counts at critical locations	☐ Vehicle classifications
Traffic Counts	Movement Counts	☐ Pedestrian volumes by crosswalk ☐ Bicycle volumes by intersection approach and movement ☐ Light vehicle volumes by intersection approach and movement	☐ Heavy truck volumes by intersection approach and movement ☐ Transit volumes by intersection approach and movement
Other Traffic Characteristics		☐ Vehicular speed ☐ Travel time	☐ Queues ☐ Vehicular delay by movement
Intersection Geometry and Traffic Control Devices		☐ Travel lanes ☐ Storage bays ☐ Pedestrian crosswalks ☐ Intersection widths ☐ Prohibited turning movements ☐ One-way streets ☐ Approach grades ☐ Sight-line restrictions ☐ On-street parking ☐ Loading zones	☐ Transit stops ☐ Profile grade ☐ Intersection skew angle ☐ Adjacent land uses ☐ Nearby access points ☐ Existing vehicular and pedestrian signal displays ☐ Control equipment ☐ Existing detectors and associated detector channels
Existing Signal Timing		☐ Phase sequence ☐ Limitations in the phase sequence due to geometric issues ☐ Use of overlaps ☐ Yellow change ☐ Red clearance ☐ Minimum green ☐ Maximum green ☐ Walk interval ☐ Flashing don't walk interval	☐ Passage time ☐ Detector settings ☐ Conditional service ☐ Time-of-day plans ☐ Coordinated phase(s) ☐ Cycle length ☐ Splits and/or force-offs ☐ Offsets ☐ Offset reference point
Crash History		☐ Crash type ☐ Crash severity	☐ Environmental collisions ☐ Causal factors

3.3.1 Field Visits

Field visits are an essential part of developing quality signal timing plans.

Observing the study area during the expected timing plan periods (e.g., weekday a.m. peak, weekday p.m. peak, and weekend midday peak) will be informative for the development of signal timing. As part of the field visit, the practitioner should

- Take photos and/or video of the intersection operations.

- Adopt the role of the different types of intersection users:
    - *Motorized Vehicle Driver:* Drive the corridor making a sample of movements (e.g., through, left turn, and minor street).
    - *Pedestrian:* Walk the intersection.
    - *Bicyclist:* Bike the intersection.
    - *Transit User:* Access nearby stations and ride transit.
- Observe the critical intersection movements, progression interactions, and user characteristics and interactions/conflicts in a stationary position over multiple cycle lengths.
- Consider upstream and downstream bottlenecks that may be influencing traffic demand.
- Think ahead to the impact of timing changes.

Particularly at intersections near capacity, minor changes in signal timing can have major impacts (both positive and negative) and should be based upon an informed decision. Multi-day observations are desirable to validate signal timing plans and their effectiveness at accomplishing planned outcomes/operational objectives.

3.3.2 Traffic Counts

Often referred to as "counts," the quantification of traffic volumes by user type is an important fundamental measure that guides signal timing development for both intersections and networks. There are generally two types of counts conducted as part of a timing project: 24-hour weekly counts and peak-period movement counts by user type. Preferably, these count data are available on an ongoing basis or, at a minimum, represent multiple time periods and days of the week sufficient to reflect system variability. However, these counts only represent a snapshot in time, so it is advantageous to continuously monitor traffic volumes at a congestion-free, mid-block location. This type of monitoring allows a practitioner to track volume changes over time, as well as determine whether 24-hour and peak-period counts are representative of seasonal and yearly trends.

Twenty-four-hour weekly counts and peak-period turning movement counts should be collected as part of signal timing development.

The practitioner should be aware that counts may be inaccurate due to data collection errors, and it is possible for traffic to vary significantly from day to day, week to week, and month to month. Too often, data are entered into a model and the results used without the practitioner verifying how slight volume changes will affect operations. For example, coordination is often run well into the evening under the assumption that late-night traffic patterns are the same as those during the peak period, when that might not be the case depending on the operating environment. It cannot be overemphasized that traffic counts are, at best, a snapshot in time, and need to be applied with judgment and understanding of the effects on signal operations.

An important limitation of counts taken at traffic signals is that they may not represent actual demand. **This is a major mistake that is very common in practice.** Traffic ***demand*** is the total number of users **desiring to reach** a certain point in a network. Traffic ***volume*** is the total number of users that **can reach** that certain point in the network. One way to measure demand is to set an intersection in free operation (with adequate maximum green times to clear standing queues) and count the volume of traffic. This technique can only be used to measure demand when there is adequate

*Traffic volumes do **not** represent traffic **demand** if queuing exists.*

upstream storage for the queue and there is no downstream congestion blocking departing traffic.

For under-capacity conditions, demand is equivalent to the measured traffic volume. However, if more vehicles arrive for a movement than can be served, the movement is considered to be operating over capacity (also called oversaturation). When an intersection movement (or movements) approaches or arrives at capacity, then additional delays occur in the network. Additional information on oversaturation is presented in Chapter 12.

Unless the practitioner measures demand arriving at the intersection through queue observation (vehicles unserved at the end of green) or measurement of departure rates from an upstream under-capacity (also called under-saturated) intersection, the true demand at an over-capacity intersection may be unknown. This can cause problems when developing signal timing plans. For example, time may be added to a given movement, only to have it used up by the previously unserved demand and possibly transfer the over-capacity problem to another location. For more information on collecting traffic volumes, see the Federal Highway Administration *Traffic Monitoring Guide* and *Traffic Monitoring Guide Supplement* (*3, 4*).

3.3.2.1 Twenty-Four-Hour Weekly Counts

Weekly traffic volumes should be collected at critical locations along the study corridor using temporary road tubes, system detectors, or detectors at signalized intersections in place along the corridor. Using existing system detectors or intersection detectors will reduce the time and cost of the data collection effort, but the sources should be validated in order to understand the accuracy of the resulting data.

Volume profiles developed from the weekly counts are an important element in the data collection effort and are used to identify

- The number of timing plans that should be used during the weekdays and weekend,
- When to transition from one timing plan to the next,
- Directional distribution of traffic along the corridor, and
- Times to collect peak-period turning-movement counts.

Exhibit 3-12 illustrates how traffic volumes can change over a 24-hour period. In this case, traffic peaks in the morning and evening. There are two cycle lengths that have been selected at this location—a longer cycle length of 110 seconds during the peak periods (7:00 a.m. to 10:00 a.m. and 3:00 p.m. to 8:00 p.m.) and a shorter cycle length of 85 seconds between the peak periods (10:00 a.m. to 3:00 p.m.). Generally, it is best to start the longer cycle length before it is needed. This allows the transition to finish before the longer cycle length is required and gives the intersection time to recover from any negative impacts of transition (described in detail in Chapter 7) (*5*). Note that in the example in Exhibit 3-12, the intersection is operating fully-actuated during the late night and early morning.

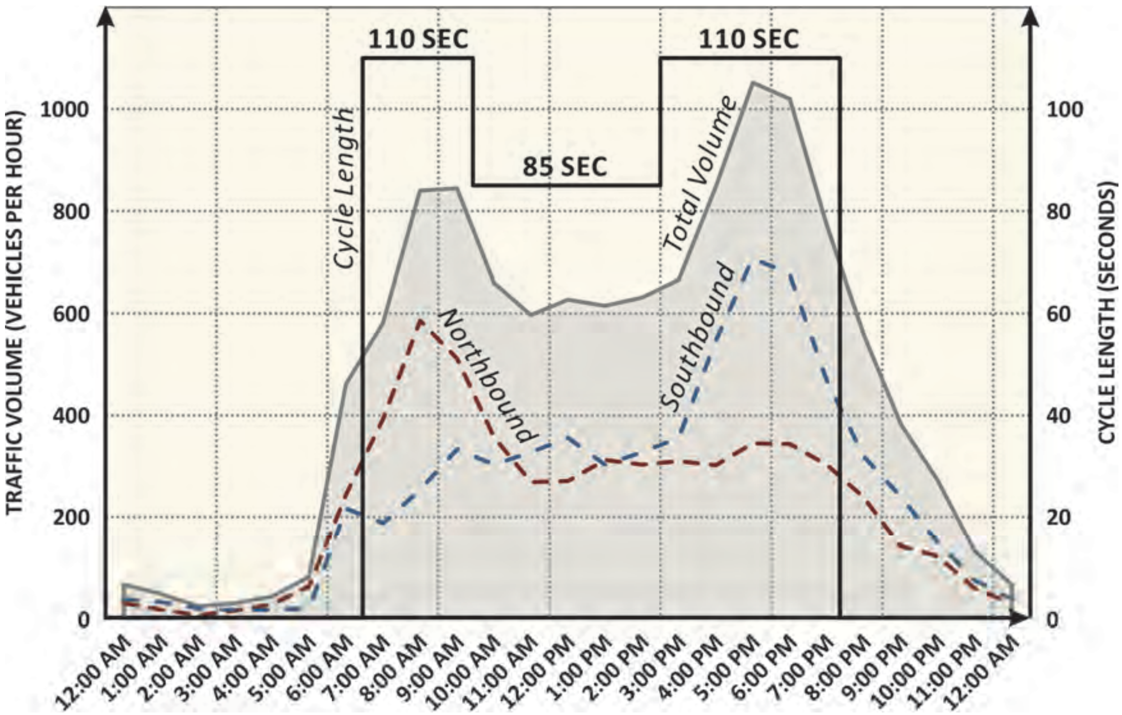


Exhibit 3-12 Example Vehicle Traffic Volume Profiles and Time-of-Day Cycle Lengths

3.3.2.2 Movement Counts

Volumes from movement counts are used to assess the existing and future operations at an intersection. Movement counts are typically collected at each study intersection during representative traffic periods, which can be identified based on the daily traffic volume profiles (discussed in the previous section). Depending on the traffic volumes and traffic patterns along the corridor, movement counts are often only conducted during peak periods, commonly the weekday morning, midday, and evening peak time periods.

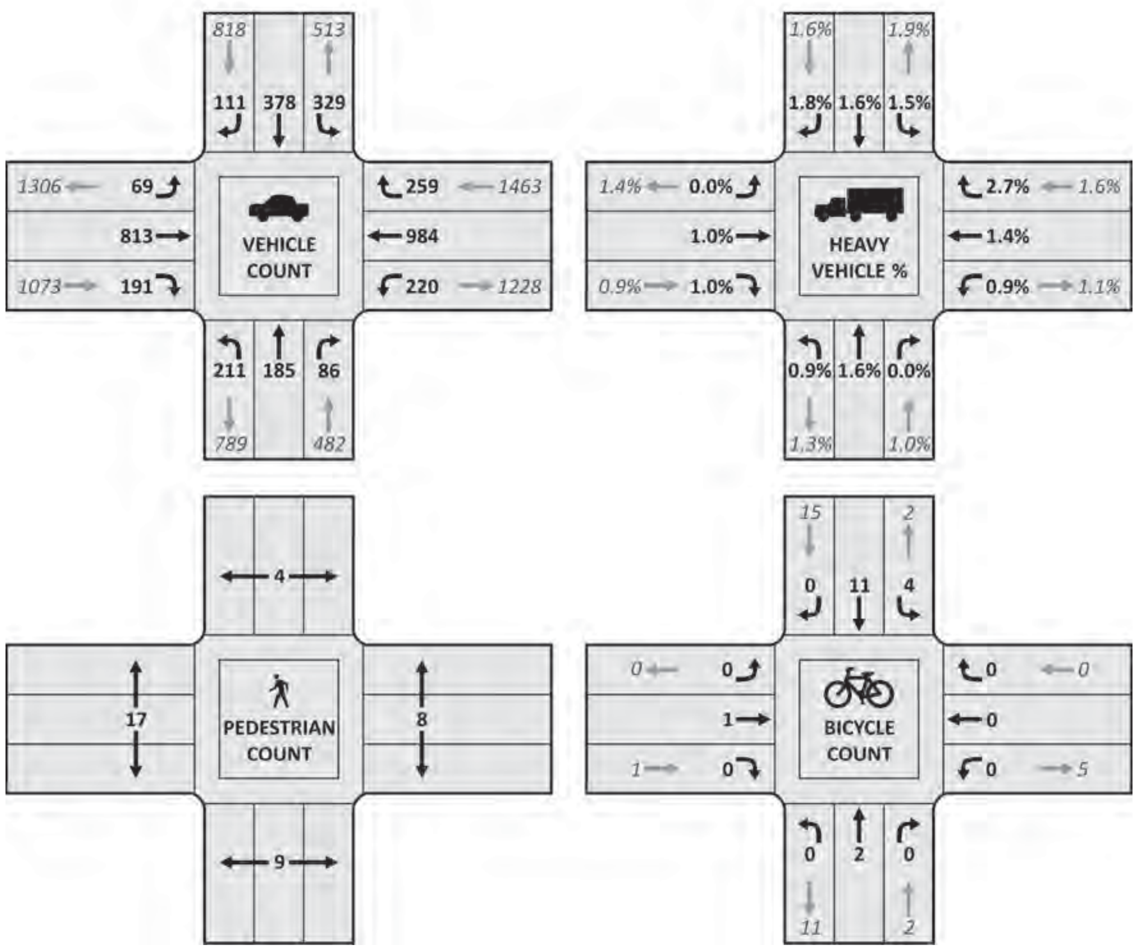
For other time periods (i.e., weekend or off-peak), the daily traffic volume profiles can be used in combination with the peak-period movement counts to develop volume adjustment factors, which can be used to estimate the volumes for any time period. At some locations, the traffic volumes may be higher or more critical during the weekend time periods, which could warrant performing movement counts for both weekday and weekend time periods. Additionally, seasonal traffic patterns may need to be considered and incorporated in the scope of work. Procedures for conducting, recording, and summarizing traffic counts are described in the *Manual of Transportation Engineering Studies* (*6*).

Exhibit 3-13 presents categories of traffic count data that are typically collected as part of movement counts as well as data that should be collected, if possible, with the understanding that not all information is needed or useful in all scenarios. In general, traffic counts should document volumes, categorized by intersection approach (e.g., northbound or southbound), movement, and mode. If heavy vehicles are a significant portion of the vehicular traffic, the traffic count may need to be further categorized by vehicle classification. The presence of special users at an intersection (e.g., elderly pedestrians, school children, or emergency vehicles) should also be documented. Exhibit 3-14 shows an example of turning-movement count data, including vehicle counts by approach and movement, heavy-vehicle percentages, pedestrian counts, and bicycle counts.

Exhibit 3-13 Traffic Count Data

Typical Traffic Count Data Collected	Additional Traffic Count Data Collected If Possible
☐ Pedestrian volumes by crosswalk ☐ Bicycle volumes by movement ☐ Light vehicle volumes (throughput, not necessarily demand) by movement ☐ Heavy-vehicle (trucks and buses) volumes by movement	☐ Queuing at beginning of first count period and end of each count period ☐ Vehicle volumes by lane (i.e., lane utilization) ☐ Transit volumes (e.g., percent buses, light rail, and streetcar) ☐ Number of pedestrian actuations (i.e., ratio of cycles with pedestrian service to total cycles) ☐ Number of preemptions and/or disrupting occurrences (e.g., trains or emergency vehicles) ☐ Number of signal priority requests and services (e.g., transit or truck) ☐ Number of slow accelerating vehicles at/near front of queue ☐ Saturation flow rate or car-following/driver behavior characteristics (qualitative or quantitative)

Exhibit 3-14 Example Traffic Volume Data



If possible when conducting movement counts (whether done in the field or using video), the presence of queues should be noted at the start of the counts and at the end of every counting period (typically every 15 minutes). If counts are performed using video, the practitioner should review the associated video to check the quality of the counts; understand the actual traffic demand if greater than the capacity; and evaluate the influences of various modes, nearby bus stops, driveways, on-street parking, and weather conditions that may affect intersection performance (see Chapter 11 for weather-related issues).

Turning-movement counts are often one of the more costly items in a traffic signal timing analysis. FHWA published *Signal Timing on a Shoestring* (*7*), which provides some guidance on ways to minimize the data collection effort through a short-count method or by estimating turning-movement counts using link traffic volume data. These guidelines may be useful on corridors or networks where the traffic conditions are predictable. However, when developing timings where traffic patterns are rapidly changing, it may be more appropriate to collect 2- to 3-hour turning-movement counts for the various study time periods.

3.3.3 Other Traffic Characteristics

Other traffic characteristics (e.g., speed, travel time, and queues), in addition to traffic volumes, may be beneficial to the signal timing plan development process. Not only will these data be useful when developing the signal timing plans, but these traffic characteristics can also be used as performance measures to evaluate the effect of signal timing changes. Performance measures are explained in detail in Section 3.4.

Some traffic characteristics can be used to guide the selection of timing values as well as be used to compare the before-and-after effect of signal timing plans.

3.3.4 Intersection Geometry and Traffic Control Devices

Intersection geometry can influence the order of movements at an intersection, as well as the amount of time given to those movements. The signal timing practitioner should take an active role in new intersection design or intersection improvements, as the best signal timing plan cannot overcome bad geometrics. A site survey should be conducted to record relevant geometric information (using a condition diagram like the example shown in Exhibit 3-15), including

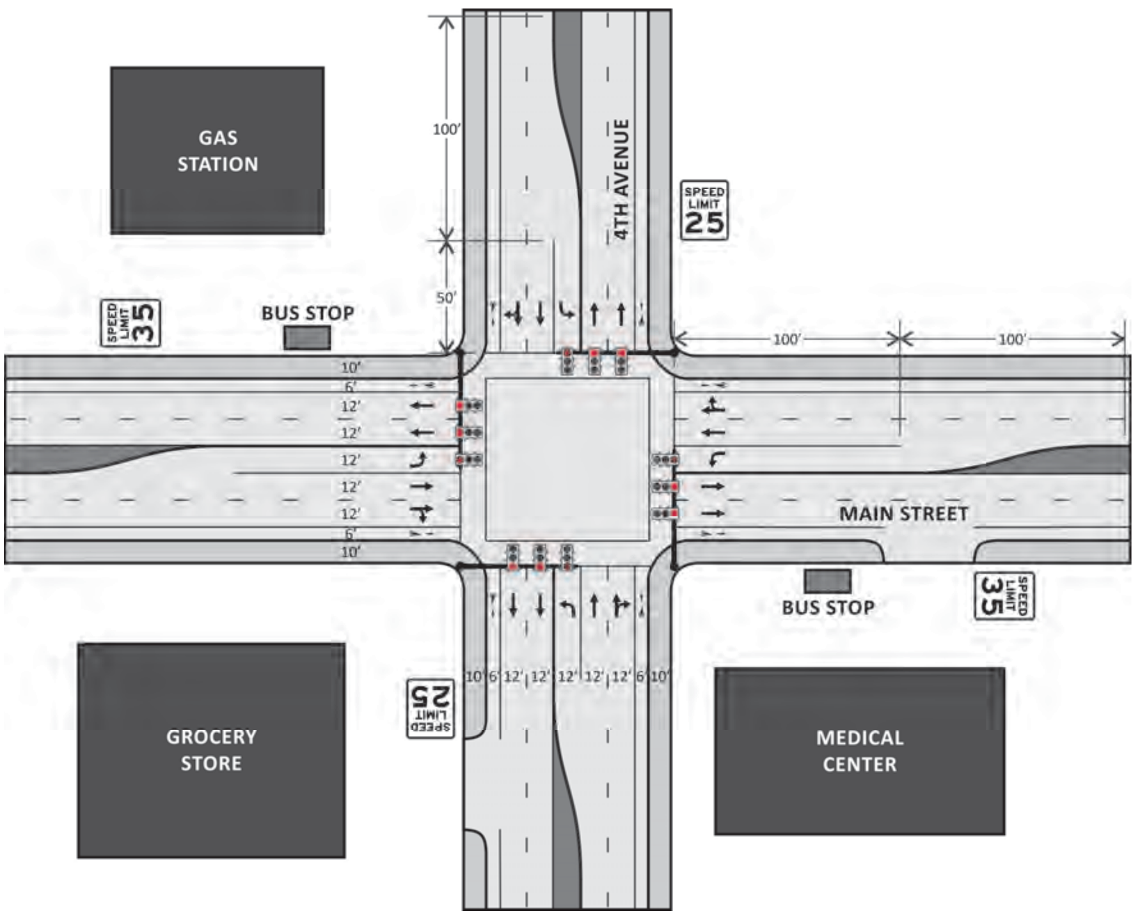
- Number of lanes (regular and special purpose, including transit and bicycle),
- Speed limits,
- Lane widths,
- Lane assignment (e.g., exclusive left turn, through only, or shared through/right turn),
- Presence of storage bays,
- Length of storage bays,
- Length of pedestrian crosswalks,
- Intersection widths for all approach legs,
- Presence of prohibited turning movements,
- Presence of one-way streets,
- Approach grades,
- Sight-line restrictions (buildings and foliage),
- Presence of on-street parking,

- Presence of loading zones,
- Presence of transit stops,
- Profile grade near the intersection,
- Intersection skew angle,
- Adjacent land uses, and
- Access points near the intersection.

It is also important to record information about the existing traffic control devices, in order to determine the capabilities, limitations, and functionality of the signal equipment. Necessary information includes

- Number, location, and type of existing vehicular, pedestrian, bicycle, and transit signal displays;
- Type of control equipment (requires agency coordination for access to cabinet);
- Number, location, length, type, and condition of existing detectors (vehicular, pedestrian, bicycle, transit, and railroad); and
- Number of detector channels associated with the detection for each traffic movement.

Exhibit 3-15 Example Intersection Condition Diagram



3.3.5 Existing Signal Timing

The existing signal timing can help the practitioner understand what currently exists in the field and provide a baseline for improvement. Key information to obtain from the existing signal timing (explained in detail throughout Chapters 5 and 6) includes

- Phase sequence (ring-and-barrier diagram),
- Limitations in the phase sequence due to geometric issues,
- Use of overlaps,
- Yellow change and red clearance intervals,
- Minimum green and maximum green,
- Pedestrian walk and flashing don't walk (FDW) intervals,
- Passage time,
- Detector settings,
- Time-of-day plans, and
- Special features.

If the intersection is operating in coordination, then the following information (explained in detail in Chapter 7) will be required in addition to that listed above:

- Coordinated phase(s),
- Cycle length,
- Splits and/or force-offs,
- Offsets, and
- Offset reference point (it is ***critical*** to understand what the controller actually uses as its offset reference, so that it can be correctly related to the optimization method used).

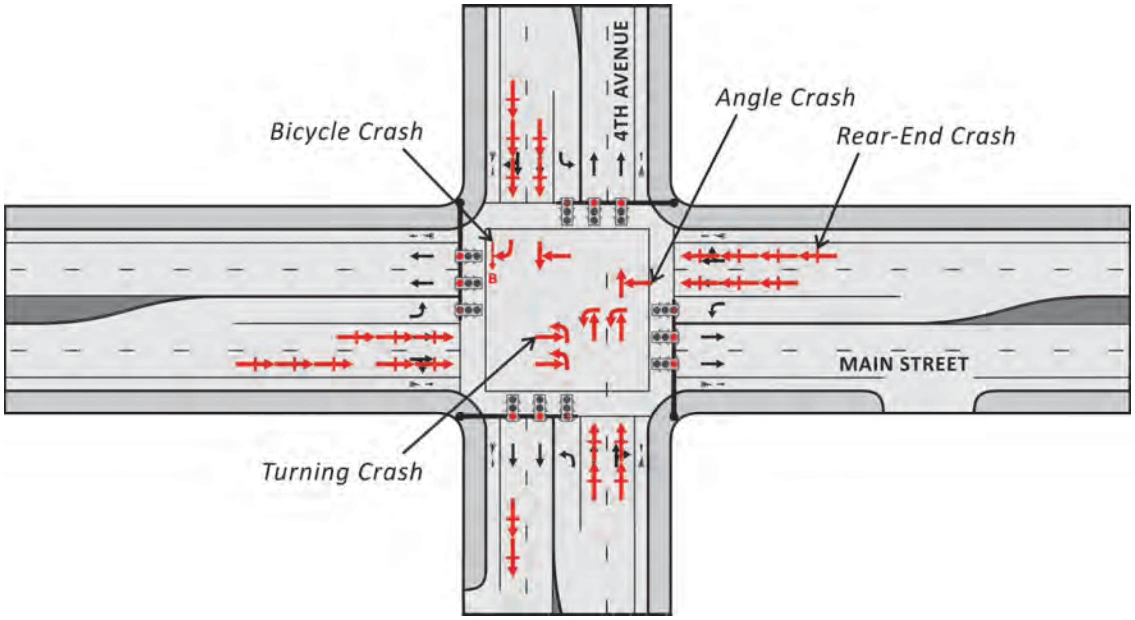
If observations are made that reveal the existing signal timing plan is working adequately, then this should be noted. Many locations have traffic patterns that may be stable and require few signal timing changes, except those related to changes in recommended practices (such as clearance intervals). The existing timing may have evolved over time in response to site-specific conditions. This initial review may result in focusing efforts on limited issues or even not undertaking timing changes, except when specific problems are identified.

If the signal timing isn't broken, don't try to fix it.

3.3.6 Crash History

Signal timing is one of the many factors that may contribute to crashes, so practitioners should be aware of the potential impacts that signal timing changes can have on safety. The practitioner should obtain 3 to 5 years of crash data; summarize the crashes by type, severity, and environmental conditions; and prepare a collision diagram of the crashes to help identify trends (as shown in Exhibit 3-16).

Exhibit 3-16 Example Signalized Intersection Crash Diagram



After reviewing historical crash data, a site visit can help the practitioner identify causal factors. For example, field observations may indicate that drivers in some movements are more aggressive than others or that certain movements are consistently running red lights. These characteristics may contribute to an increased number of crashes and could be the result of signal timing parameters (e.g., insufficient green time, poor offset, or long cycle length). The practitioner should evaluate the effect of parameters such as

- Clearance intervals;
- Detector settings, size, and locations;
- Cycle length;
- Offsets; and
- Phase sequence.

A variety of tests and other evaluation tools are discussed further in FHWA's *Signalized Intersections: An Informational Guide* (*8*).

3.4 OPERATIONAL OBJECTIVES AND PERFORMANCE MEASURES

Operational objectives and performance measures are at the heart of the outcome based process. The objectives and their associated performance measures tangibly define user priorities by movement for the location being timed. Instead of choosing signal timing values based solely on software outputs (such as vehicle delay), the outcome based process develops signal timing values based on objectives and performance measures that have been selected for a specific location. Note that some performance measures should be simple and understandable to system users, such as stops per mile and travel time. Others, such as system delay, may be helpful to system operators but of little interest to users.

3.4.1 Select Operational Objectives

By clearly establishing objectives, a practitioner will be able to choose timing values that reflect user needs. Objectives may be chosen based on known problems in the study area, as a result of public comments, staff observations, or known discrepancies with established policies. The identification of problems during timing analysis may be an iterative exercise done in conjunction with determining objectives. In other words, previous problems may help shape the operational objectives and/or, by clearly defining operational objectives, deficiencies may be more apparent and easily addressed.

Objectives will help focus the goals for a signal timing project.

There are many possible objectives for traffic signal operations, and signal timing strategies will change depending on the objective(s) that are chosen. The practitioner should understand that some objectives may be mutually exclusive. For example, objectives that focus on pedestrians will require different signal timing solutions than objectives that focus on vehicles. Below is a list of 15 specific operational objectives that can be used individually or in combination to focus signal timing efforts. Not all are useful in conveying performance in an understandable manner to the traveling public.

3.4.1.1 Vehicle-Specific Objectives

1. ***Vehicle Safety*:** Minimize vehicle collisions, reduce vehicle conflicts, and provide sufficient time for vehicles to execute movements.
2. ***Vehicle Mobility—Capacity Allocation*:** Serve vehicle movements as efficiently as possible, while also distributing capacity as fairly as possible across movements and modes. Prioritize movements according to need without excessively delaying other movements.
3. ***Vehicle Mobility—Corridor Progression*:** Move vehicles along high-priority paths (typically along high-volume movements on corridors) as efficiently as possible without excessively delaying other movements.
4. ***Vehicle Mobility—Secondary Progression*:** Move vehicles along low-priority paths (such as major street turning movements or minor street movements) as efficiently as possible without excessively delaying other movements.
5. ***Environmental Impact Mitigation*:** Minimize the amount of pollution induced by improving the efficiency of vehicle trajectories, either by reducing delay or stops or using infrastructure-to-vehicle technology to influence vehicle travel patterns.
6. ***Queue Length Management*:** Prevent formation of excessive queues on critical lane groups, such as freeway exit ramps.
7. ***Vehicle and Driver Costs*:** Minimize stops and delay in order to reduce vehicle operating costs and driver delay costs.

3.4.1.2 Pedestrian-Specific Objectives

1. ***Pedestrian Safety*:** Minimize pedestrian involvement in collisions, reduce pedestrian conflicts, and provide sufficient time for pedestrians to execute movements.
2. ***Pedestrian Mobility*:** Serve pedestrian movements as efficiently as possible.

3. ***Pedestrian Accessibility*:** Provide the ability for pedestrians, including special-needs groups, to execute movements.

3.4.1.3 Bicycle-Specific Objectives

1. ***Bicycle Safety*:** Minimize bicycle involvement in collisions, reduce bicycle conflicts, and provide sufficient time for bicycles to execute movements.
2. ***Bicycle Mobility*:** Serve bicycle movements as efficiently as possible.

3.4.1.4 Transit-Specific Objectives

1. ***Transit Safety*:** Minimize transit vehicle involvement in collisions, reduce transit vehicle conflicts, and provide sufficient time for transit vehicles to execute movements.
2. ***Transit Mobility*:** Serve transit vehicles as efficiently as possible.
3. ***Transit Accessibility*:** Provide the ability for transit vehicles to execute maneuvers (i.e., loading/unloading activities) and for transit users, including special-needs groups, to access transit.

3.4.2 Establish Performance Measures

Performance measures (also known as measures of effectiveness or MOEs) should be selected for each operational objective to evaluate the success of the timing plan(s). Most practitioners use delay and stops as measures because they are readily calculated in optimization software. However, optimization software will not always calculate performance measures that are appropriate for the operational objectives. For example, some models optimize ***system*** delay, which is not perceived by users, while others combine performance measures (e.g., performance index that combines stops and delay).

Users perceive stops first and ***their*** delay second. Minimizing stops on arterials has the most impact on reducing user complaints, and this objective might be achieved through strategies that increase system delay slightly. For example, longer cycle lengths may be selected to reduce arterial stops, which could also result in an increase to system delay. In the field, arterial travel time studies have been the chief means of assessing performance.

Some objectives (such as pedestrian priority) may need to be assessed qualitatively, as the performance measures may be difficult to measure in the field or through most analysis tools. However, objectives should not be overlooked simply because they are not available through software applications. Some signal timing decisions may require a qualitative assessment. Exhibit 3-17 presents a variety of traditional vehicle performance measures that can be collected to measure signal timing effectiveness. While these are the performance measures that are most often used and easily measured or estimated, they should only be used when they are consistent with the desired outcomes. Information about non-motorized performance measures is available in Section 3.4.3.

In addition to summarizing timing parameters and data collection methods for each vehicle performance measure, Exhibit 3-17 also indicates (1) whether each performance measure is generally used to evaluate an individual intersection or system of intersections and (2) whether it is typically measured in the field or derived from a software program or related measure. The performance measures are described in detail below.

Exhibit 3-17
Performance Measures, Timing Parameters, and Collection Methods

Objective	Performance Measure	Intersection	System	Measured	Derived	Influential Signal Timing Parameter(s)	Typical Method of Collection
Corridor Progression	Quality of Progression – Percent Arrival on Green	X		X	X	□ Cycle length □ Offset □ Phase order	□ Accurate detection and controller logging
	Quality of Progression – Ratio of Arrival on Green to Arrival on Red		X	X		□ Cycle length □ Offset □ Phase order	□ Probe vehicle (GPS floating car, GPS fleet data)
	Number of Stops per Mile		X	X		□ Cycle length □ Split □ Offset □ Phase order	□ Probe vehicle (GPS floating car, GPS fleet data)
	Travel Time/ Average Speed		X	X		□ Cycle length □ Split □ Offset	□ Probe vehicle (GPS floating car, GPS fleet data, Bluetooth™, or other re-identification)
Capacity Allocation	Delay (Vehicle or Person)	X		X	X	□ Cycle length □ Split □ Offset □ Phase order	□ Typically estimated from models □ Manual queue estimation method □ Call to service controller logging
	Phase Failures	X		X	X	□ Cycle length □ Split □ Phase order	□ Queue detection plus signal indication data (manual observation or automated technology) □ Approximate with phase termination logging (max outs/force-offs versus gap outs)
	Queuing	X		X	X	□ Cycle length □ Split □ Offset □ Phase order	□ Manual observation or automated technology (video, radar) □ Approximate with occupancy
Vehicle Safety	Safety-Related	X		X		□ Detection location and settings □ Clearance intervals	□ Rate of deceleration from vehicles □ Red-light running (cameras, controller logging) □ Number of conflicts □ Crash records
Combination	Composite Index	X	X		X	□ Cycle length □ Split □ Offset	□ Equation that combines multiple performance measures

Practitioners should reference the FHWA *Traffic Analysis Toolbox* (*9*) and *Performance Measurement Fundamentals* (*10*) for more information on data collection practices. Field data collection may provide the most accurate performance measures, but can often be cost or resource prohibitive. Many performance measures can be approximated through the use of traffic operations/safety software. The use of software can provide a lower-cost approximation of a variety of performance measures. However, some field data are often necessary or helpful in creating models that reflect reality and to ensure a level of validation and reasonableness in the results where possible. Chapter 5 contains more information about software considerations.

3.4.2.1 Quality of Progression—Percent Arrival on Green

Percent arrival on green is a measure of the proportion of users (typically vehicles) that arrive during a green (or walk) indication relative to those that arrive during a red (or don't walk) indication at a particular intersection. This is best measured in the field, which can be done with some modern controller software. Optimization software programs are often used to approximate the effectiveness of progression between intersections, typically through the use of a time-space diagram (explained in detail in Chapter 5). However, field observations may also be useful when attempting to determine the quality of progression. In order to maximize the accuracy of percent arrival on green accounted for in the field, detection should be located beyond queued vehicles, yet as close to the intersection stop bar as practical.

3.4.2.2 Quality of Progression—Ratio of Arrival on Green to Arrival on Red

A similar performance measure to percent arrival on green is the ratio of intersections that a user (typically a vehicle) arrives at on green (or walk) versus arrives at on red (or don't walk). This type of performance measure compares the quality of progression between different corridors. Even corridors with the best coordination will have vehicles that arrive on a red indication at some intersections, and this performance measure provides an easy, corridor-level analysis of progression quality.

3.4.2.3 Number of Stops per Mile

Although the ratio of arrival on green to arrival on red provides a convenient way to compare multiple corridors, a high value for this ratio may be hard to achieve if a corridor does not have many signalized intersections. The number of stops per mile may provide a better snapshot of the condition of the corridor operations. It is a performance measure that is based on the relationship among user arrivals at an intersection, the phase status, and the queue profile. As noted previously, users are more aware of stops than they are of minor differences in delay.

Additionally, if improved air quality or reduced emissions are desired objectives, then minimizing the number of stops may be an important performance measure. Research has documented the negative impacts associated with deceleration and acceleration of vehicles with respect to emissions and air quality.

3.4.2.4 Travel Time/Average Speed

Travel time studies evaluate the overall quality of traffic movements along an arterial. Travel time is an easily understood and intuitive performance measure to the general public, operators, planners, and maintenance staff. Practitioners should refer to the FHWA *Travel Time Data Collection Handbook* (*11*) and *Travel Time Reliability Measures* (*12*) for additional information on travel time data collection.

While travel time provides an easily understood overall performance measure, it does not reveal the underlying traveler experience, which is sensitive to stops first and delay (especially small amounts that may be imperceptible) second. Use of travel time should not be done at the expense of ignoring stops and delay (two outcomes of the selection of cycle length, offsets, and splits). Modern controllers can monitor arrival on green and estimate approach delay, allowing practitioners to adjust timing parameters and measure results at the intersection level, while also looking at overall travel time as an aggregate measure of performance.

3.4.2.5 Delay (Vehicle or Person)

Delay is the difference in travel time that a user experiences between free-flow (unimpeded) conditions and current conditions. It is a primary measure in optimization models because it is easily quantified. It can also be used in models to estimate users' operating costs. However, incremental changes in delay at an intersection are less noticeable to roadway users than other mobility-related performance measures, such as number of stops or overall trip travel time. It is also not readily measured in the field.

Delay at a signalized intersection can be the result of (1) signal control and timing, (2) queues that impede travel, or (3) factors such as bus blockages, parking maneuvers, and distracted drivers. Understanding if and to what degree these forms of delay contribute to user experiences can be important and should be noted during field observations. Delay can ultimately be expressed in two ways:

1. Unit delay (seconds/vehicle), which is related to the user's perception of disutility at an intersection; or
2. Total accumulated delay (vehicle-hours), which is related more to the economic performance of an intersection. One vehicle-hour of delay is accumulated when one vehicle is delayed for a full hour, or 3600 vehicles are delayed for 1 second each, etc. (*13*).

3.4.2.6 Phase Failures

A phase failure (also called a "split failure" when the intersection is coordinated) is defined as the occurrence of one or more stopped vehicles that cannot proceed through a signalized intersection on a green indication. Occasionally, a movement may be served twice per cycle, in which case the term "cycle failure" may be used if all vehicles are not served within a cycle. Phase, split, or cycle failures are typically easy to observe in the field; the challenge for the practitioner is to determine how or if correction is necessary.

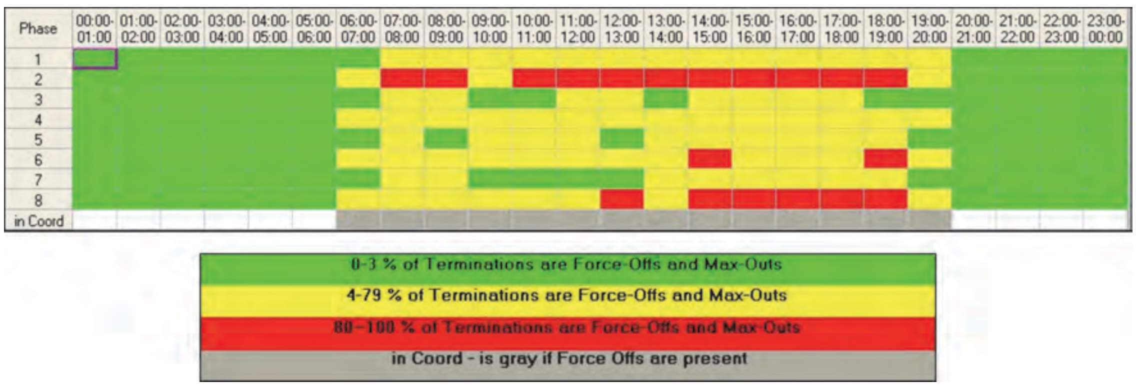
A simplified approach for approximating phase failures in an automated fashion is measuring and recording signal phase termination types. This involves comparing the proportion of gap outs (which imply enough time to serve demand) to max outs or force-offs (which imply not enough time to serve demand). Because the coordinated phases are typically not actuated, information may not be available for those phases.

Exhibit 3-18 illustrates an example of individual and aggregated gap out versus max out/force-off conditions at a signal. Visualization graphics of gap out versus max out/force-off conditions are also available (depicted in Exhibit 3-19) to aid in identifying problematic movements. Reviewing this type of information is emphasized in NCHRP Project 3-79, “Measuring and Predicting the Performance of Automobile Traffic on Urban Streets” (*14*).

Exhibit 3-18 Example Split Termination Logging

681	Record Number	216				Sample Period			15
682	Date	Wednesday, February 09, 2011				Mid - Period Time			17:07
683	Phase	1	2	3	4	5	6	7	8
684	Phase Service	8	7	4	6	8	7	6	8
685	Ped Service	0	1	0	1	0	1	0	2
686	Average Green	18	44	20	39	20	42	8	32
687	Max Outs	0	0	0	0	1	0	0	0
688	Force Offs	4	6	0	4	2	5	2	8
689	Gap Outs	4	1	4	2	5	2	4	0

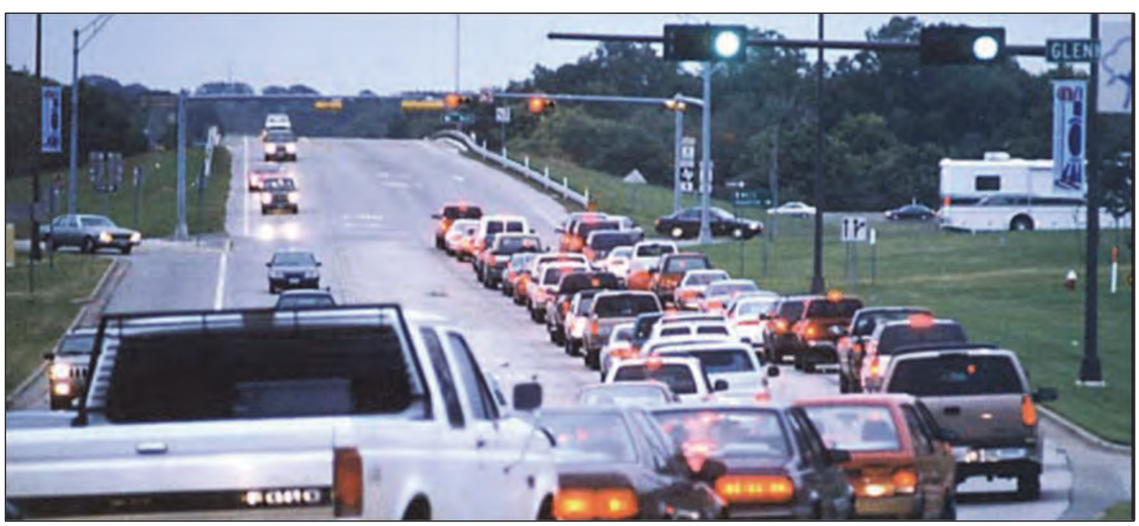
Exhibit 3-19 Example Split Termination Logging Visualization



3.4.2.7 Queuing

Queuing is the length or number of users waiting for a green (or walk) indication. It is typically measured at the beginning of the green (or walk) indication when the standing queue length is longest. Queuing is often a direct result of the signal timing and detection parameters selected, but it is important to recognize that queues can also be influenced by downstream or upstream traffic operations. For example, in Exhibit 3-20, green time is being wasted at the nearest signal due to a queue and red-light display at the downstream signal.

Exhibit 3-20 Arterial Efficiency Decreased by Signal Timing



Queues can be estimated using optimization software, but conducting a field visit is the best way to evaluate conditions. A field visit, both before and after implementation of a new timing plan, provides the opportunity to observe operational issues such as storage bay blocking or spillback (depicted in Exhibit 3-21 and Exhibit 3-22, respectively) and approaches that are not serving the full demand. One of the challenges of field data collection is visiting the corridor in a way that allows observation of the most critical time periods and important intersections.

Exhibit 3-21 Storage Bay Blocking

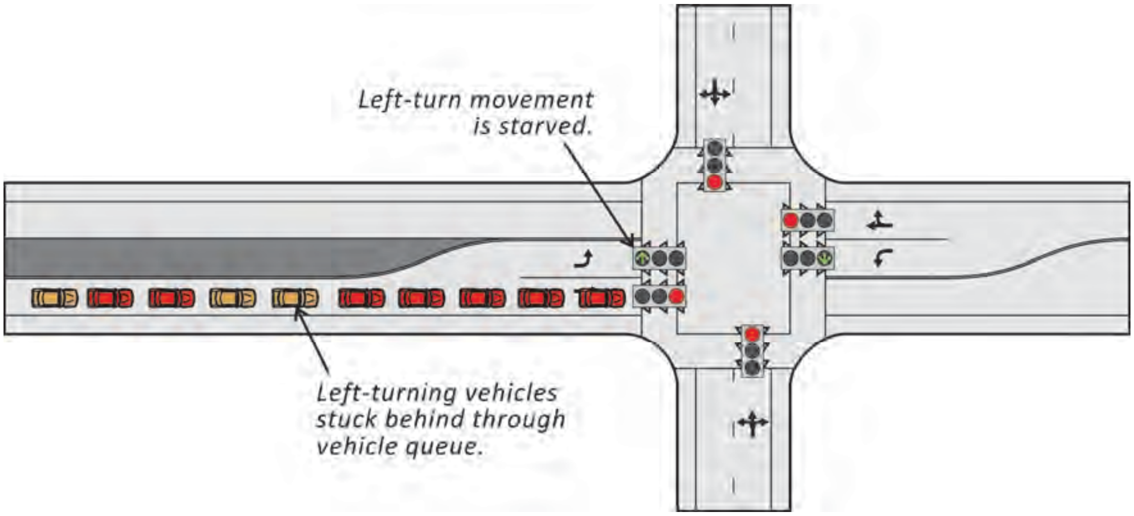
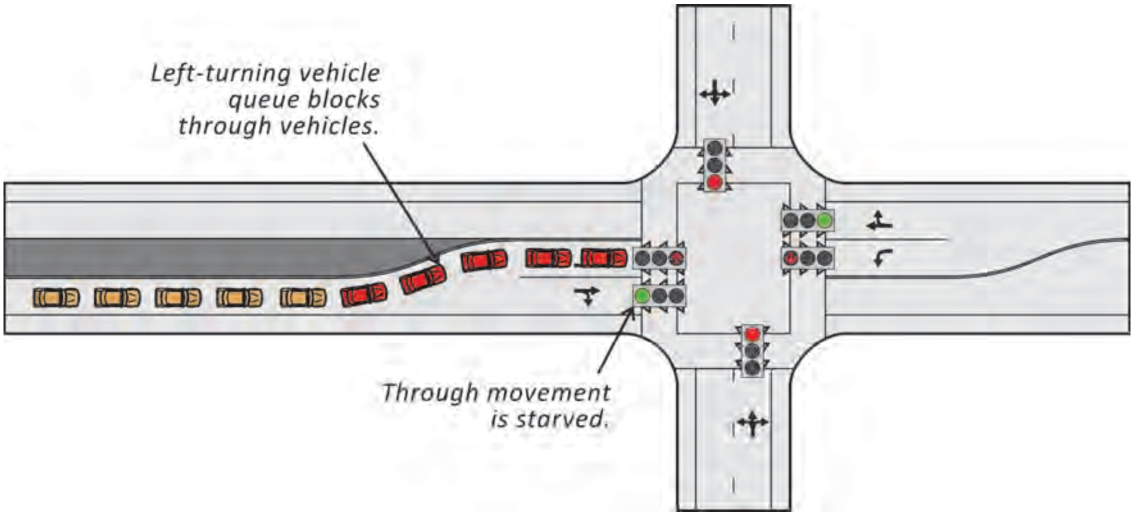


Exhibit 3-22 Storage Bay Spillback



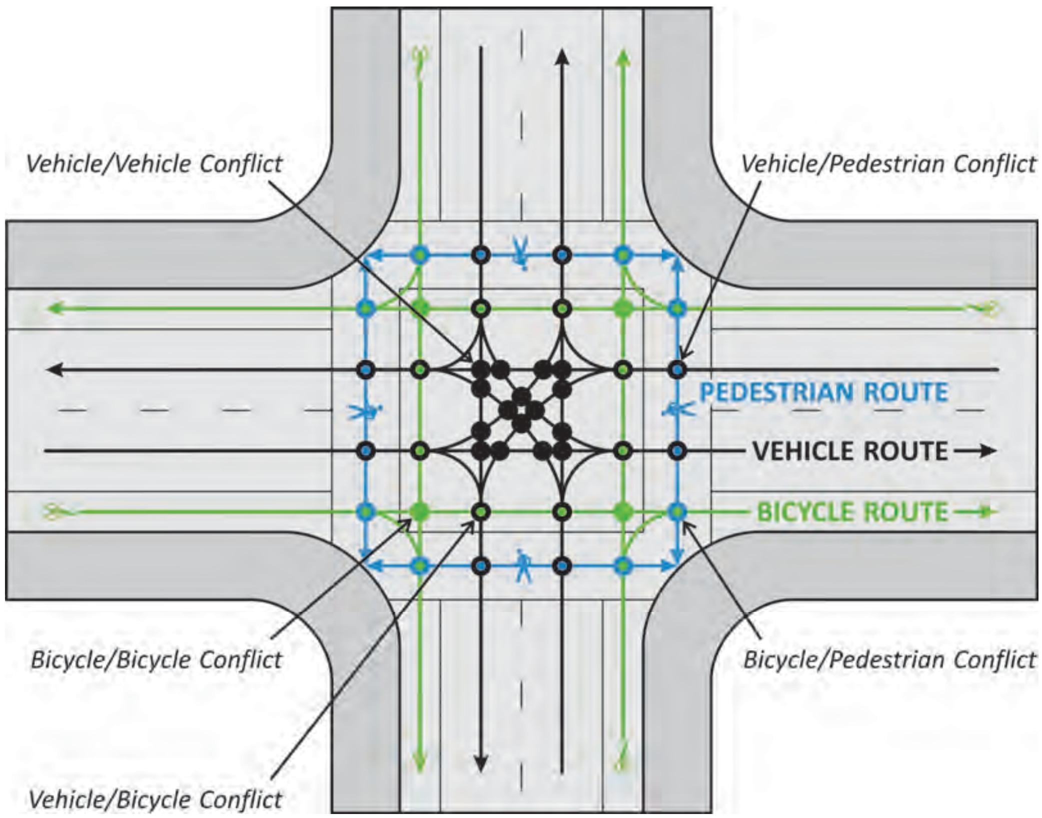
3.4.2.8 Safety-Related Performance Measures

The following bullets summarize common safety-related performance measures that are influenced by signal timing. Additional information on safety-related performance measures can be found in the *Highway Safety Manual* (*15*).

- ***Number, type, and severity of crashes*** can be influenced by signal timing parameters for both multiple-vehicle and single-vehicle collisions. Types of crashes at a signalized intersection can include angle, head-on, rear-end, sideswipe, and collisions with pedestrians and bicycles. Severity ranges from property-damage-only to various levels of injury to fatal. However, it should be noted that signal timing is only one factor that influences crashes, even at a signalized intersection. Crash reports are often available through the jurisdiction that operates and maintains the intersection.

- ***Number of potential conflict points***, which are locations where two movements conflict at an intersection. Exhibit 3-23 illustrates all of the potential conflict points at a standard signalized intersection with permitted left-turn phasing. The number of conflict points can be influenced by the type of phasing that is chosen. However, phasing can also influence conflict points outside of the intersection (i.e., turn-lane conflicts), so should be chosen carefully.

Exhibit 3-23 Conflicts between Movements at an Intersection



- ***Frequency of pedestrian conflicts*** with vehicles and bicycles can influence safety for many modes at an intersection. Pedestrian visibility to turning vehicles and bicycles is an important qualitative performance measure.
- ***Frequency of bicycle conflicts*** with conflicting vehicle movements can be another qualitative performance measure. Conflicts can be reduced by providing sufficient time for bicycles to start moving and proceed through an intersection before conflicting movements.
- ***Frequency of max out or force-off conditions due to vehicular demand*** can be a safety-related metric as the signal controller will terminate a phase once the max out or force-off is reached, regardless of vehicle proximity to the stop bar or intersection.
- ***Frequency of users violating the red (or solid don't walk) indication*** can be an important metric for understanding where and for what reason(s) red-light

violations may be occurring (as shown in Exhibit 3-24). This metric can be tracked through manual observation or automated technologies, including video or any detection zone that can identify a vehicle entering the intersection on red.

Exhibit 3-24 Red-Light Violation

3.4.2.9 Composite Index

One performance measure may not accurately represent the operation at a signalized intersection or along a corridor. Some agencies have considered using an "index" that combines two or more performance measures. This composite allows agencies to include multiple performance measures in their assessment. For example, the Orange County Transportation Authority *Regional Traffic Signal Synchronization Master Plan* (*16*) explains how to combine three performance measures—average speed, green signals per red signals, and stops per mile—into a Corridor Synchronization Performance Index (CSPI). This index is an easy way to compare multiple corridors using several performance measures.

3.4.3 Non-Vehicle Operational Objectives and Performance Measures

The practitioner should remember that not all operational objectives are easily measured. In some cases, they may be qualitative in nature. Objectives should ***not*** be chosen because they are easy to measure. They should be chosen after the desired outcome is understood. Exhibit 3-25 is a matrix that summarizes some non-vehicle performance measures as they relate to operational objectives. While this is not meant to be an exhaustive list, it is meant to demonstrate additional multimodal objectives and potential performance measures that can be chosen to evaluate a desired outcome. Just as the practitioner should carefully choose the operational objectives based on a specific operating environment, he or she must also carefully choose performance measures to serve as true representations of the outcome based objectives.

Exhibit 3-25 Non-Vehicle Operational Objectives and Performance Measures

Operational Objective		Percent Arrival on Green (or Walk)	Ratio of Arrival on Green (or Walk) to Arrival on Red (or Don't Walk)	Number of Stops per Mile	Travel Time/ Average Speed	Delay	Phase Failures	Queuing	Accident Rate	Composite Performance Measure
		Performance Measure								
Pedestrian	Safety								X	X
	Mobility	X	X		X	X				X
Bicycle	Safety								X	X
	Mobility	X	X	X	X	X				X
Transit	Safety								X	X
	Mobility	X	X	X	X	X	X	X		X

3.5 REFERENCES

1. *A Policy on Geometric Design of Highways and Streets,* 6th Edition. American Association of State Highway and Transportation Officials, Washington, D.C., 2011.
2. *Manual on Uniform Traffic Control Devices for Streets and Highways,* 2009 Edition. United States Department of Transportation, Federal Highway Administration, Washington, D.C., 2009.
3. *Traffic Monitoring Guide*. Office of Highway Policy Information, Federal Highway Administration, United States Department of Transportation, Washington, D.C., 2001.
4. *Traffic Monitoring Guide Supplement*. Office of Highway Policy Information, Federal Highway Administration, United States Department of Transportation, Washington, D.C., 2008.
5. Shelby, S. G., D. M. Bullock, and D. Gettman. Resonant Cycles in Traffic Signal Control. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1925*, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 215–226.
6. Robertson, H. D., J. E. Hummer, and D. C. Nelson. *Manual of Transportation Engineering Studies*. Institute of Transportation Engineers, Washington, D.C., 2010.
7. Henry, R. D. *Signal Timing on a Shoestring*. Report FHWA-HOP-07-006, Federal Highway Administration, United States Department of Transportation, 2005.
8. Chandler, B. E., M. C. Myers, J. E. Atkinson, T. E. Bryer, R. Retting, J. Smithline, J. T. P. Wojtkiewicz, G. B. Thomas, S. P. Venglar, S. Sunkari, B. J. Malone, and P. Izadpanah. *Signalized Intersections Informational Guide*, Second Edition. Report FHWA-SA-13-027, Federal Highway Administration, United States Department of Transportation, 2013.
9. Federal Highway Administration. *Traffic Analysis Toolbox*. http://ops.fhwa.dot.gov/trafficanalysistools/. Accessed Dec. 22, 2013.

10. Federal Highway Administration. *Performance Measurement Fundamentals*. http://ops.fhwa.dot.gov/perf_measurement/fundamentals/. Accessed Dec. 22, 2013.

11. Turner, S. M., W. L. Eisele, R. J. Benz, and D. J. Holdener. *Travel Time Data Collection Handbook*. Report FHWA-PL-98-035, Federal Highway Administration, United States Department of Transportation, 1998.

12. Federal Highway Administration. *Travel Time Reliability Measures*. http://ops.fhwa.dot.gov/perf_measurement/reliability_measures/index.htm. Accessed Dec. 22, 2013.

13. *Florida Intersection Design Guide*. Florida Department of Transportation, 2013.

14. Bonneson, J. A., M. P. Pratt, and M. A. Vandehey. NCHRP Project 3-79, "Measuring and Predicting the Performance of Automobile Traffic on Urban Streets." Transportation Research Board of the National Academies, Washington, D.C., 2008.

15. *Highway Safety Manual*. American Association of State Highway and Transportation Officials, Washington, D.C., 2010.

16. *Regional Traffic Signal Synchronization Master Plan*. Orange County Transportation Authority, 2009.

CHAPTER 4
SIGNAL DESIGN

CONTENTS

LIST OF EXHIBITS

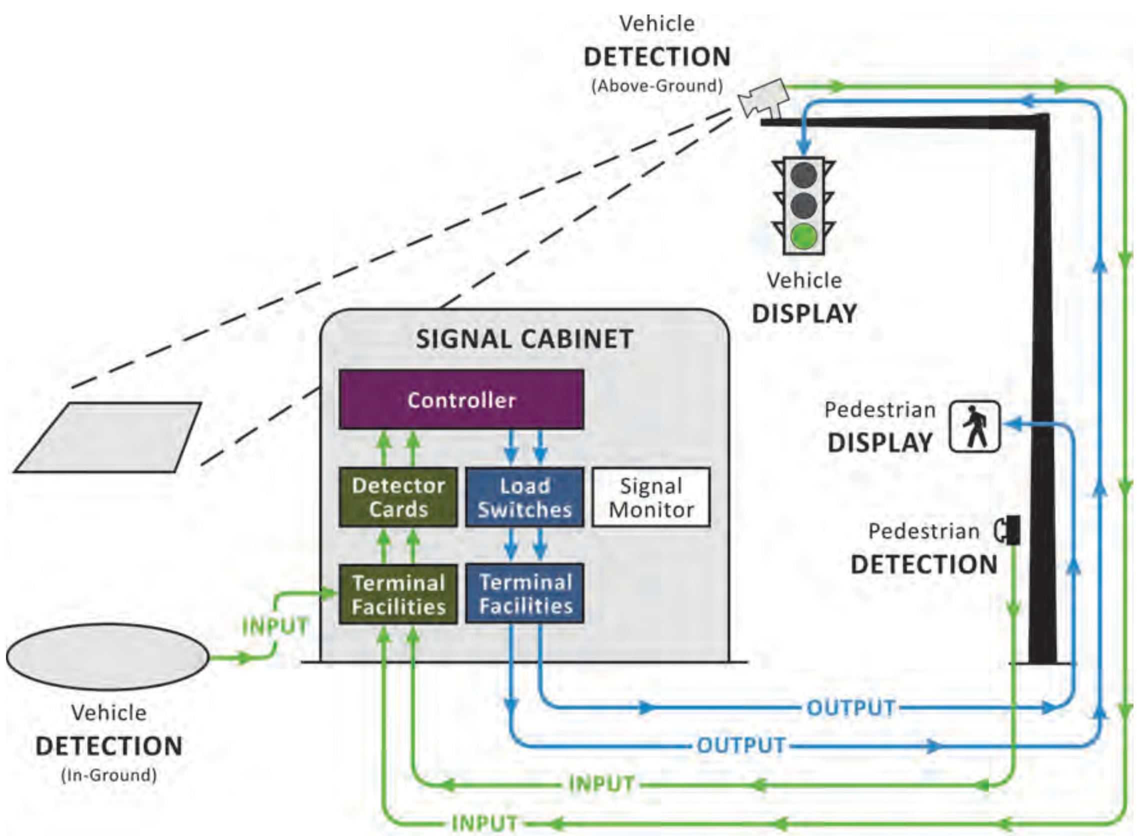
CHAPTER 4. SIGNAL DESIGN

This chapter discusses the signal design elements that directly influence signal timing, but not every feature described in this chapter will be required in every operating environment.

Effective signal timing requires appropriate signal design. While signal timing parameters can be changed relatively easily, signal design elements are typically static and more difficult (and costly) to change. In order to produce a signal timing plan that meets the objectives chosen for a site, the signalized intersection design must be tailored to its specific operating environment and work as a foundation for the signal timing. Appropriate infrastructure provides a signal timing practitioner with options, allowing him or her to adjust timing values to meet site-specific needs in the near-term as well as the future. Within agencies, signal design is typically separate from signal timing. It is therefore important for signal timing practitioners to understand the effects of signal design on intersection operations, so that they can offer guidance to signal designers and ensure all aspects of the signal timing can be accommodated.

Exhibit 4-1 illustrates the flow of inputs and outputs to and from the various pieces of signal equipment at an individual signalized intersection. As previously discussed in Chapter 3, there are three main categories of signal equipment at an intersection: (1) detection, (2) cabinet equipment, and (3) displays. Detectors sense users and provide the cabinet equipment with information about their presence at an intersection. The cabinet equipment (the controller in particular) interprets the inputs from the detectors and translates them to outputs based on the signal timing plan. The output information is then forwarded to the displays in order to allow users to move through the intersection.

Exhibit 4-1 Flow of Inputs and Outputs among Detectors, Signal Cabinet Equipment, and Displays



Using the signal equipment categories, this chapter has been organized into four main sections (as depicted in Exhibit 4-2). The first three sections summarize design information for the equipment at an individual signalized intersection, while the last section focuses on the equipment required for system operations (i.e., communications for monitoring and coordination). In addition to the information provided in this chapter, a signal designer should consult the current edition of the *Manual on Uniform Traffic Control Devices* (MUTCD) (*1*) and any design standards relevant to the local jurisdiction before completing a signalized intersection design.

Exhibit 4-2 Chapter Organization Based on Signal Equipment

4.1 DETECTION

Phases are a way for the controller to time multiple movements based on the desired outcome. (Detailed information is available in Chapter 5, but phases will be referenced at a high level throughout the rest of this chapter.)

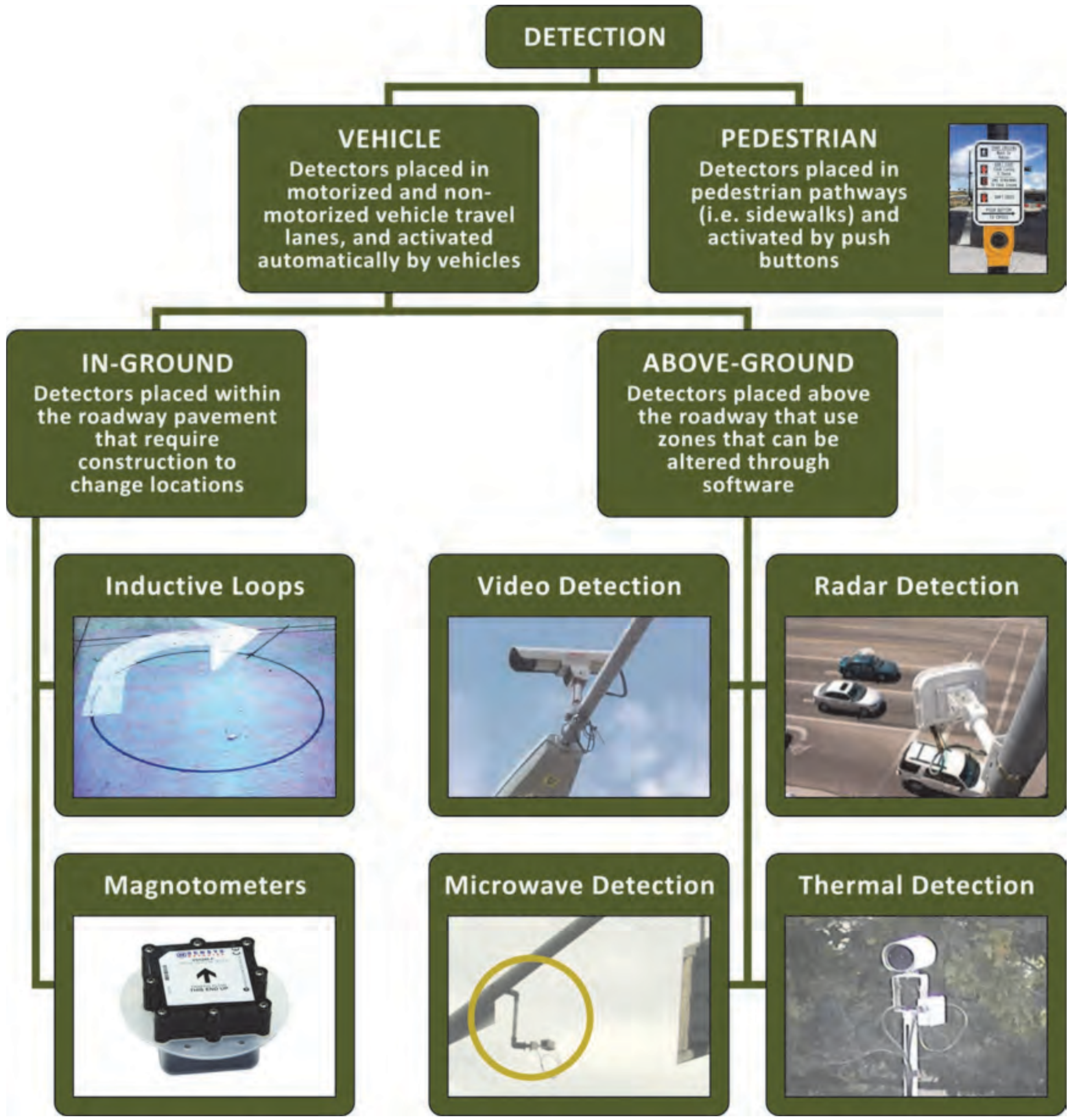
Detectors sense the presence of roadway users, and provide the controller with information it can use to determine whether a particular phase needs to be served. Detection should be designed with an understanding of actual controller operations because detectors can operate in one of two modes—pulse or presence. This manual assumes detection operates in presence mode because it is the safest mode for traffic signal operations.

The detection layout should change depending on how the detectors are being operated and will affect a variety of intersection operations. In combination with the controller, detectors are responsible for

1. Identifying user presence for a movement and its corresponding phase.
2. Extending a phase.
3. Identifying gaps in traffic where a phase should end due to no traffic or inefficient flow.
4. Providing safe phase termination for high-speed vehicle movements by minimizing the chance of a driver being in the decision zone (also known as a Type II dilemma zone) at the onset of the yellow indication (explained further in Section 4.1.1).
5. Monitoring intersection performance using measures of effectiveness (MOEs) logs.
6. Possibly counting traffic volumes and identifying vehicle types (e.g., trucks, bicycles, emergency vehicles, and transit vehicles). Chapter 10 contains additional information on special applications.

Exhibit 4-3 shows several categories of detectors that can be used at an intersection. While the same type of detection can be used to detect motorized vehicles and bicycles, push buttons located near the sidewalk ramps are most appropriate for pedestrians. Motorized vehicle and bicycle detectors can either be in-ground (units placed in the roadway pavement) or above-ground (units positioned above the roadway on signal mast arms or light poles).

Exhibit 4-3 Types of Traffic Signal Detection



Designers should first refer to adopted local or state practices, guidelines, or policies, if available, when selecting a detection technology. Practitioners can also refer to the Federal Highway Administration's *Traffic Detector Handbook* (*2*) for additional information about these different types of detection. Each has advantages and disadvantages related to performance, reliability, and installation. While most designs will utilize the same type of detection along a single corridor, these detector types can be used in combination.

When developing detection plans, designers need to consider the potential signal phasing options, expected travel speeds, and user mix at the intersection (e.g., number of trucks and slow-moving pedestrians). An agency should always consider maintenance efforts associated with the installation of new equipment, but this chapter explains concepts assuming detection on all approaches. Traffic signal detection on all approaches provides the most flexibility, allowing for effective responsiveness and facilitating the measurement of demand and travel characteristics over time. Detection on all approaches also allows adaptable late-night operations.

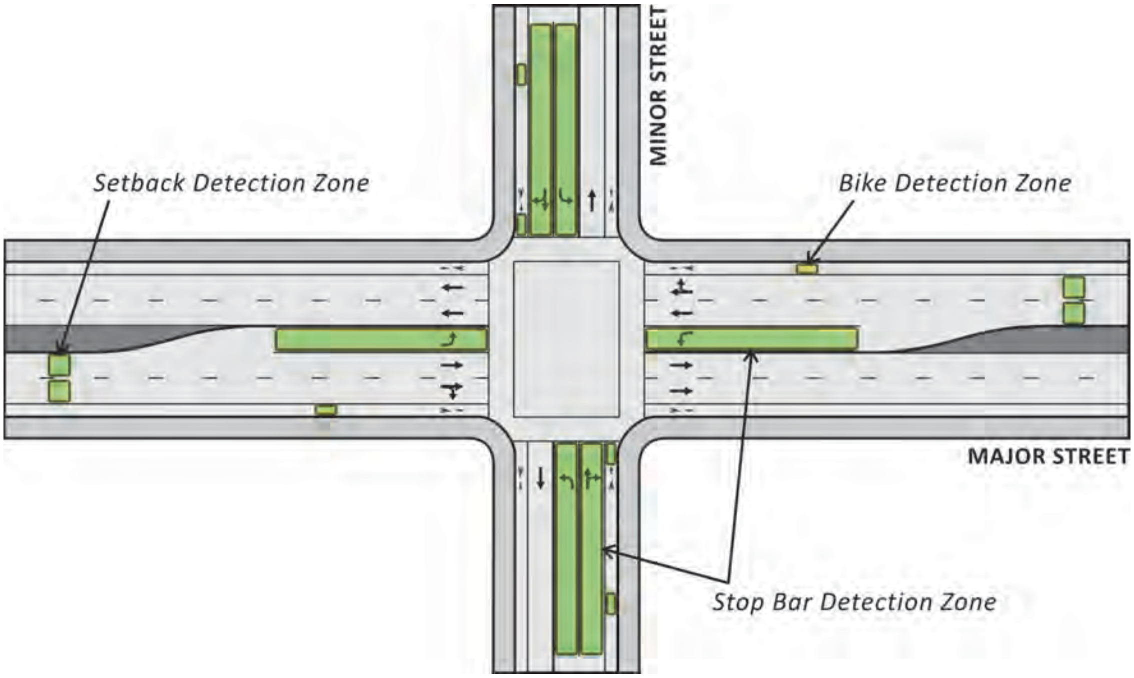
Generally accepted vehicle detection locations, in order of priority, are as follows:

1. Minor street approaches and left-turn lanes at the stop bar.
2. Major street approaches set back from the stop bar.
3. Major street approaches at the stop bar.

Depending on the objectives for the signalized intersection, detection may only be required at some locations. Each detection location has its own benefits. Stop bar presence detectors are able to drop calls when permitted turns are made (i.e., left-turn or right-turn-on-red) and no other vehicles are present, reducing inefficient transitions. Setback detectors, on the other hand, can provide decision zone (Type II dilemma zone) protection and have the ability to provide more efficient gap outs than is possible with stop bar detection zones. (Gapping out is explained in detail in Chapter 6.) Combinations of setback detection and stop bar detection can be used in either through lanes or left-turn lanes to increase efficiency.

Detection zones can vary based on location-specific issues, the type of detection technology used, and the approach speed. Exhibit 4-4 illustrates a basic approach for detection zone placement at the intersection of a major street, high-speed approach (shown left/right) and a minor street, low-speed approach (shown top/bottom). Note that the stop bar detection is shown as long zones. Depending on the type of detection technology, these zones can be made up of a single longer zone or multiple smaller zones. For example, if inductive loops are being used, a stop bar detection zone might be made up of three inductive loops.

Exhibit 4-4 Basic Detection Layout



Detection zones can also be tied together, meaning that several detection zones can essentially send the same message to the controller. For example, if there are two detectors in separate lanes tied together, a vehicle can drive over either one of them to register a call for that movement. This is inefficient from a timing standpoint. For detection zones being used to extend (or call and extend) a phase, detection channels should be designed by lane for more efficient lane-by-lane detection and the ability to count users on a lane-by-lane basis.

In addition to detector locations, detection plans often include a description of the size, number, and functionality of each detector, as well as a wiring diagram that shows how detectors are associated with phases. In general, detectors can either call and/or extend a phase. If a phase is not being served and vehicles approach the intersection, the detectors can tell the controller that there are vehicles waiting and "call" the phase. If a phase is in the process of being served and additional vehicles approach the intersection after the initial queue departs, the detectors can ask the controller to "extend" the current phase to accommodate the vehicles. Exhibit 4-5 summarizes the primary objectives for different types of detection, with detailed information provided throughout the remainder of Section 4.1.

Exhibit 4-5 Detection Objectives

Type of Detection	Primary Objective(s)
High-Speed-Approach Vehicle Detection	□ Serving the standing queue at the beginning of green □ Safely terminating the phase once there is a conflicting call
Low-Speed-Approach Vehicle Detection	□ Calling phases on low-speed approaches □ Serving the standing queue at the beginning of green □ Minimizing delay by reducing calls on permitted movements
Left-Turn-Lane Vehicle Detection	□ Calling left-turn phases □ Serving the standing queue at the beginning of green □ Minimizing delay by reducing inefficient transitions □ Preventing vehicles from being stranded in the intersection
Right-Turn-Lane Vehicle Detection	□ Minimizing delay by reducing calls due to right-turn-on-red □ Calling right-turn phases (if used)
Pedestrian Detection	□ Calling pedestrian phases
Bicycle Detection	□ Calling either associated motorized vehicle phases or independent bicycle phases □ Preventing accidental motorized vehicle actuations if using independent bicycle phases □ Eliminating need for bicycles to use pedestrian phases
Emergency Vehicle Detection	□ Enabling preferential treatment options for emergency vehicles
Bus Detection	□ Enabling preferential treatment options for transit
Rail Detection	□ Ensuring safe and efficient signal timing sequencing before, during, and after train arrivals

4.1.1 High-Speed-Approach Vehicle Detection

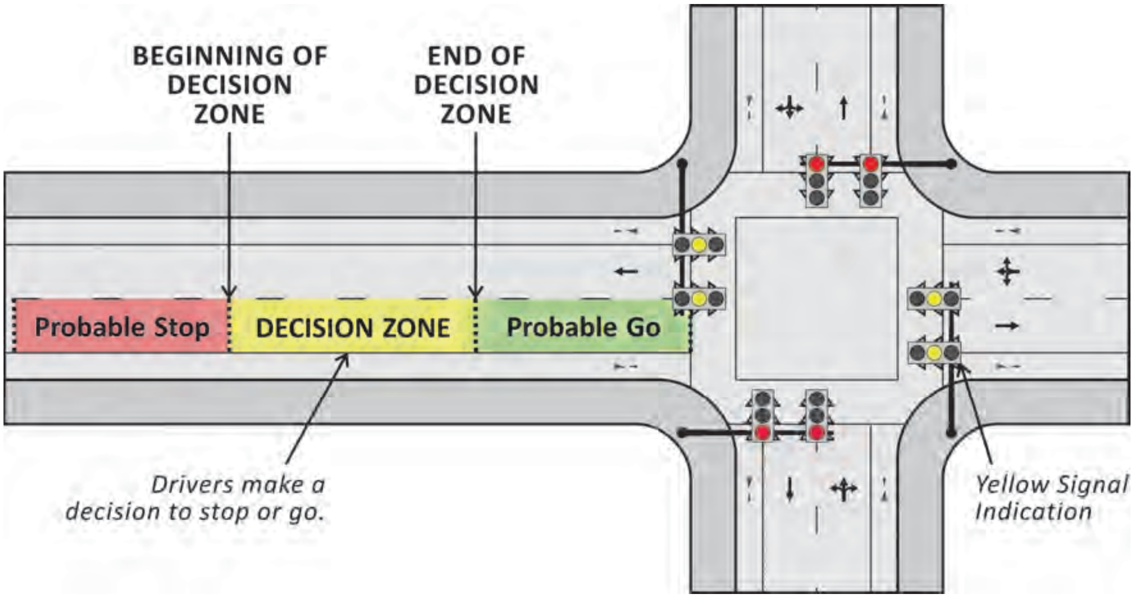
High-speed detection requires detectors located upstream of the stop bar in order to determine when it is safe to terminate a phase as well as for measuring queue and performance. While setback (or advance/upstream) detection can also be used to clear queues at the stop bar (using minimum green or variable initial, which are discussed in Chapter 6), stop bar detectors can be used on high-speed approaches in combination with setback detection. Stop bar detectors can more efficiently clear queues at the stop bar and can accommodate calls if there is traffic that enters the roadway downstream of the setback detectors. Therefore, a combination of stop bar and setback detection provides the most efficient operations. If used with setback detection, stop bar detectors should be disconnected (using controller programming) once the initial queue clears.

This will prevent the phase from being extended unnecessarily and allow for more efficient operations.

The design of setback detection on high-speed approaches requires special attention. The detectors must be placed in a manner that allows the controller to terminate phases while vehicles still have enough time to stop. This relationship is often described using the term "dilemma zone," which has historically been applied to two scenarios that are often confused. The term dilemma zone was initially used with regard to yellow change intervals (i.e., Type I dilemma). Later, the term dilemma zone was used with regard to detection design (and also known as an indecision zone, Type II dilemma zone, or decision zone). This manual uses the term ***dilemma zone*** with regard to yellow change interval timing and ***decision zone*** with regard to setback detection design.

Dilemma zones are a result of the yellow clearance interval timing—when a yellow is too short for a vehicle to safely enter the intersection, but the vehicle is too close to the stop bar to safely stop. A dilemma zone can be addressed with an appropriate yellow clearance interval, which is explained in detail in Chapter 6. A decision zone is not related to clearance interval timing, but rather to the human factors of driver perception, reaction, and judgment. It is the length of roadway where each individual driver may make a different decision upon seeing the yellow signal indication; some vehicles may stop and others may go. The location of the decision zone is illustrated in Exhibit 4-6. For more information on Type I and Type II dilemma zones, practitioners should refer to *NCHRP Report 731* (*3*).

Exhibit 4-6 Decision Zone



The decision zone has historically been defined using a variety of measures, including distance to the stop bar (*4*, *5*), travel time to the stop bar (*6*), and stopping sight distance (*7*). Based on trends from these previous studies, the limits of the decision zone tend to be between 5.5 and 2.5 seconds of travel time from the stop bar. Exhibit 4-7 provides quick reference distances representing the beginning and end of the decision zone (if 5.5 seconds from the stop bar is considered the beginning and 2.5 seconds from the stop bar is considered the end). In order to design with the decision zone in mind, one detector (or more in some complex designs) should be placed upstream of the stop bar, starting at the beginning of the decision zone. Detection at the beginning of the decision zone can then be programmed to prevent a phase from terminating before a vehicle clears the decision zone (using the passage timer, which is discussed further in Chapter 6).

Exhibit 4-7 Limits of Decision Zone

Approach Vehicular Speed (Miles Per Hour)	Beginning of Decision Zone (5.5 Seconds from Stop Bar)	End of Decision Zone (2.5 Seconds from Stop Bar)
35	285 feet	125 feet
40	325 feet	145 feet
45	365 feet	165 feet
50	405 feet	180 feet
55	445 feet	200 feet

When a phase maxes out (as explained in Chapter 6) or is forced off by coordination (which is discussed in Chapter 7), there is no decision zone protection. The more max out (or force-off) occurrences, the less effective the decision zone protection. Conversely, the fewer max out (or force-off) occurrences, the more effective the decision zone protection. Strategies to reduce max out occurrences include lane-by-lane detection; establishing detector settings to get vehicles to the end of the decision zone; and complex, multiple-detector designs. Designers should refer to the *ITE Manual of Traffic Detector Design* (*8*) for more information.

4.1.2 Low-Speed-Approach Vehicle Detection

The objectives for low-speed approach detection are primarily calling the low-speed approach phases, clearing the standing queue, and minimizing delay. Because of the lower speeds, there is less need for decision zone protection. As shown in Exhibit 4-4, the use of stop bar detection only for low-speed approaches is typical practice. This large area detection design at the stop bar facilitates the primary objective of clearing queues without prematurely ending the phase. The use of large (80-foot) detection zones (either through a single larger detection zone or multiple smaller detectors) allows the passage time to be reduced to zero (creating more efficiency), prevents premature termination of the phase due to sluggish traffic, and allows for immediate termination when the last vehicle passes the stop bar (*9*). This detection design also allows the controller to drop a call if a vehicle turns left on a permitted movement or right-on-red (through the use of non-locking memory, which is described in Chapter 6).

4.1.3 Left-Turn-Lane Vehicle Detection

As shown in Exhibit 4-4, the recommended detection design for left-turn movements should match the design for low-speed approaches. This type of design is particularly helpful when the signal is operating in protected-permitted mode (described in detail in Section 4.3.1), as the call will be dropped if the left-turning vehicle turns during the permitted interval. If the left-turn movement operates in permitted or protected-permitted mode, it may be desirable to extend the stop bar detection zone beyond the stop bar into the intersection or crosswalk. This minimizes the potential for stranding a turning vehicle in the intersection. Note that if the left-turn movement occurs at a higher speed, then additional setback detection may be required; some designs use setback detection in left-turn lanes to increase efficiency.

4.1.4 Right-Turn-Lane Vehicle Detection

If a right-turn movement is in an exclusive right-turn lane (or lanes), detection can be provided for those vehicles. However, it may be desirable to place a delay on those detection zones. Delay will prevent a call from immediately going to the controller (explained in detail in Chapter 6). Instead, a call will be placed only if a vehicle has to wait for a certain amount of time. This prevents the phase from being prematurely called when a vehicle is able to quickly turn right-on-red.

4.1.5 Pedestrian Detection

Pedestrians are vulnerable users of the transportation system and are frequently at risk for conflicts with other users. Pedestrian detection should be placed in a clear and expected fashion to support ease of use and compliance with the Americans with Disabilities Act (ADA) requirements and the guidelines contained in the MUTCD (*1*). FHWA's *Pedestrian Facilities User Guide—Providing Safety and Mobility* (*10*) provides additional information.

Pedestrian crossings are generally provided between pedestrian generators/destinations and on all quadrants of a signalized intersection, unless a specific issue or objective would dictate otherwise. For example, a signal designer might reconsider a crossing at a location with very low pedestrian activity and very high left-turn (or right-turn) vehicular volume under protected-permitted operations. There may also be cases in which vehicular modes are explicitly prioritized above pedestrians based on local policy or practice.

If possible, wiring and the cabinet equipment should support independent operation of each crosswalk. Independent operation of each pedestrian crossing allows the total crossing time to be different between phase pairs. For example, the pedestrian crossings on the north and south sides of an intersection can be given different walk and pedestrian clearance times if they are operated independently.

Per MUTCD guidance, (1) the length of the crosswalk and (2) the distance from the pedestrian detector to the far side of the traveled way both impact the minimum required pedestrian service interval (walk and clearance times) (*1*). The longer the crosswalk or the further the distance of the pedestrian detector to the edge of the pavement, the greater the amount of time required to serve pedestrians. While the length of the crosswalk is not generally decided by the signal designer (or signal timer), he or she should attempt to locate pedestrian detection in an intuitive location, near the curb/ramp at the beginning of the painted crosswalk, in order to minimize the time required for pedestrian movements.

4.1.6 Bicycle Detection

If bicycle detection is provided (as shown in Exhibit 4-8), defining its desired functionality (e.g., call, extend, or count) will influence the size and location of the required detection. Regardless of the specific location, the signal designer may want to consider setback bicycle detection to minimize accidental vehicle actuations, which can be common when bicycle detectors are located at stop bars. This is particularly important when there are independent bicycle phases being used at an intersection.

Exhibit 4-8 Bicycle Detection Examples

4.1.7 Emergency Vehicle Detection

Emergency vehicles, depending on desired outcomes, may receive preferential treatment (priority or preemption) at signalized intersections (explained in detail in Chapter 10). Design components to consider are priority/preemption receivers at the intersection, a controller interface in the signal cabinet to call the appropriate phase(s), and a controller (hardware and firmware) to support priority or preemption service.

4.1.8 Bus Detection

Specialized bus signal phasing or timing requires bus detection to call bus priority phase(s). Bus priority adjusts signal timing while maintaining coordination (if present) as the bus approaches a signalized intersection. More information about preferential treatment, bus priority, and design considerations is available in Chapter 10.

4.1.9 Rail Detection

The MUTCD provides traffic control guidance with respect to railroads and light rail transit (LRT) (*1*). Rail that is on-street or has crossings near traffic signals has its own unique impacts on signalized traffic systems, based on its proximity to the signalized intersections and the modal priority hierarchy. Railroad and commuter rail (and some LRT) often operate in a semi-exclusive right-of-way, and will have the highest priority at a nearby signalized intersection, typically preempting signal timing. Appropriate detection for railroad vehicles (advance and simultaneous) and appropriate connections to nearby traffic signals can be important for ensuring safe and efficient signal timing sequencing before, during, and after train arrivals. Details about rail preemption are discussed in Chapter 10.

4.1.10 Other Detection

Special detection may be employed at unique locations such as movable bridges, one-lane bridges and tunnels, and metered freeway entrance ramps in order to provide preemption or priority control of the traffic signals (see Chapter 10).

4.2 SIGNAL CABINET EQUIPMENT

*Signal cabinet components should be designed for existing and **future** needs.*

The signal cabinet houses the control equipment at an individual intersection. Exhibit 4-9 illustrates and labels the typical components found inside a Type 332 signal cabinet, and Exhibit 4-10 does the same for a NEMA TS-2 signal cabinet. (Other common cabinet styles include Type 336, NEMA TS-1, and ITS.) Different cabinet styles have different layouts, internal operations, and terminology, but the general functionality and many components are largely interchangeable among modern signal hardware.

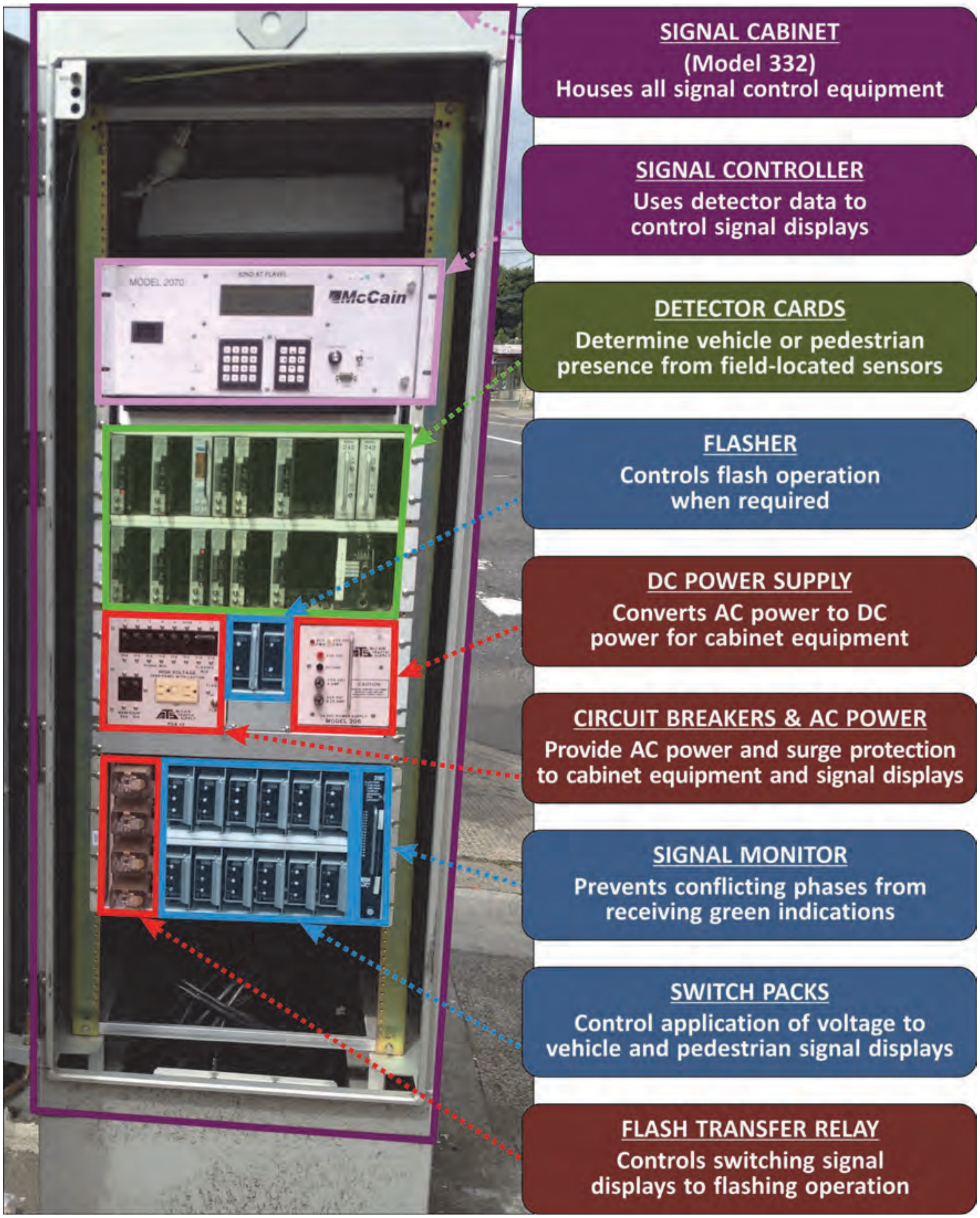
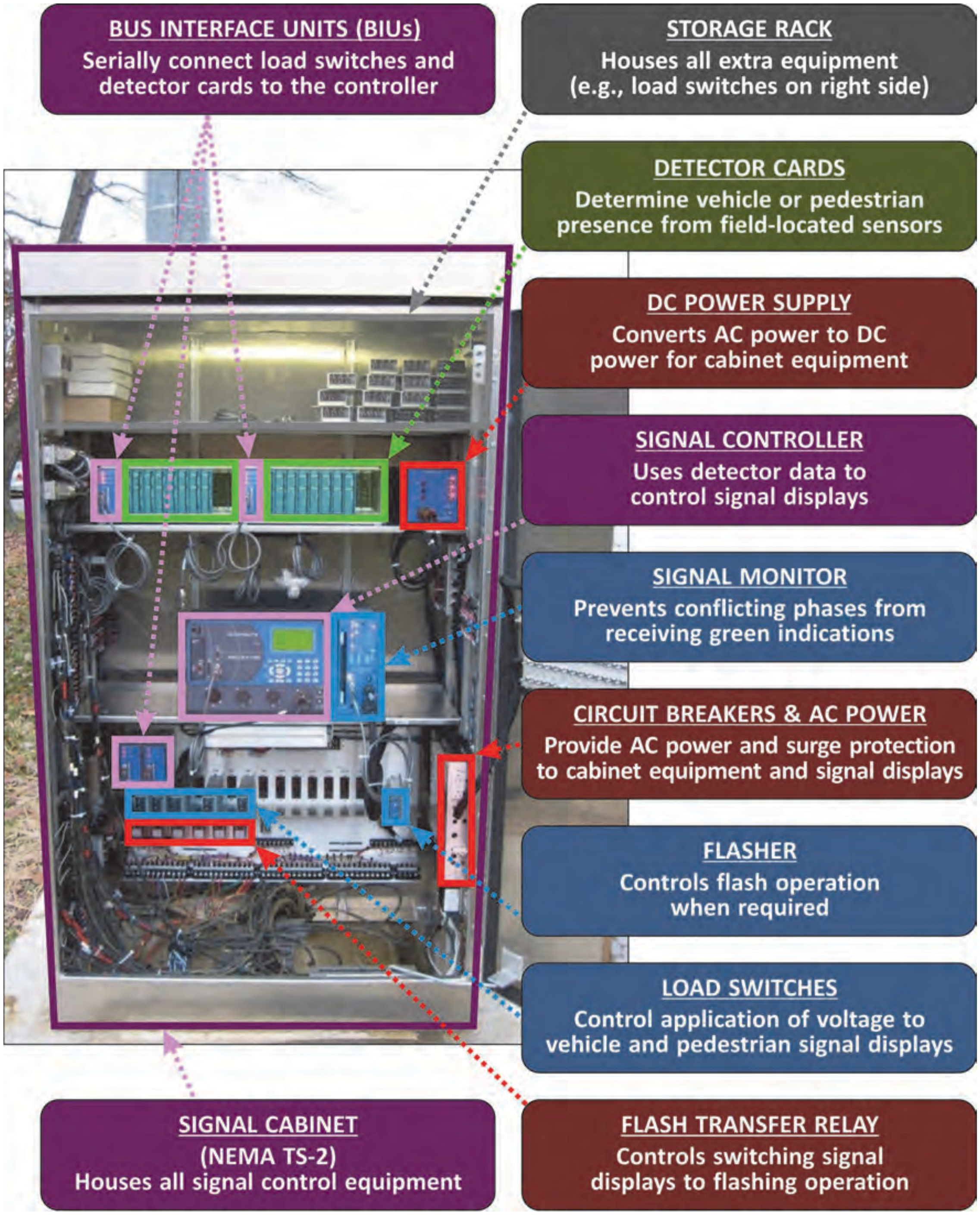


Exhibit 4-9 Basic Equipment in a Type 332 Signal Cabinet

Cabinet styles can be differentiated using several metrics (e.g., size or mounting type), but one of the most important distinguishing characteristics is how the controller is connected to other cabinet equipment. There are two types of connections—parallel and serial. A signal cabinet with parallel connections directly wires the controller to every individual input and output, while a cabinet with serial connections connects the controller to cabinet equipment through serial interface units (e.g., Bus interface units in NEMA TS-2 cabinets). Type 332, Type 336, and NEMA TS-1 signal cabinets use parallel connections, while NEMA TS-2 and ITS signal cabinets use serial connections.

Exhibit 4-10 Basic Equipment in a NEMA TS-2 Signal Cabinet



The following sections highlight a few typical cabinet components and their relationships to signal timing, but not every piece of equipment that could be in a signal cabinet is discussed in this section. Signal designers should consult relevant standard drawings and specifications.

4.2.1 Cabinet

The cabinet should be located along the minor street in a position where it is easy for a practitioner to simultaneously see the inside of the cabinet and displays for several phases, making troubleshooting and field observations more effective. Signal cabinets should also be located in a space that is easy to access (e.g., parking for maintenance vehicles), but shielded (or away) from the roadway to minimize the likelihood of crashes or “knockdowns.”

4.2.2 Controller

The controller is the piece of equipment in the signal cabinet that translates input information from the detectors into output information for the displays. Signal timing parameters (programmed into the controller software) determine how the controller interprets the detector and display information. Much like a computer, there are various components of a controller that practitioners often reference. Exhibit 4-11 relates the basic controller elements to well-known computer equivalents. In general, there is a cabinet that houses the controller, the controller that acts much like a computer, and an operating system that runs a chosen firmware.

A controller can be thought of as the “computer” behind traffic signal control.

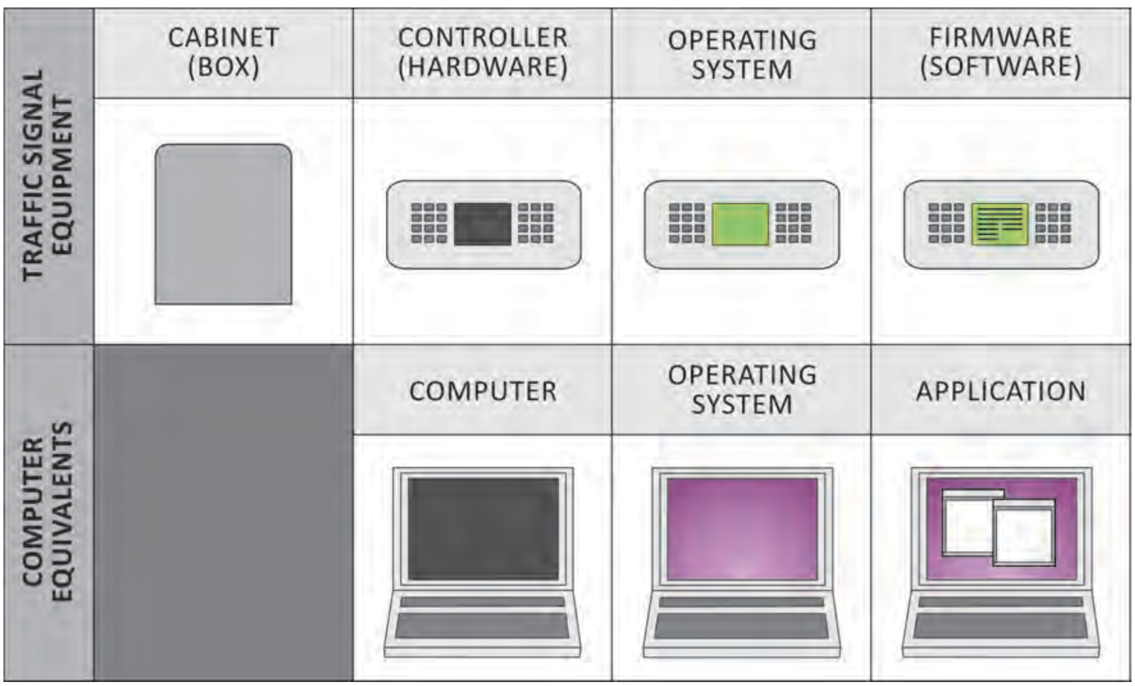


Exhibit 4-11 Signal Equipment Equivalents

Historically, there are two families of standards that have driven traffic signal control: (1) the National Electric Manufacturers Association (NEMA) family and (2) the Type 170 family, which is based on the California Department of Transportation (Caltrans) standard. While controllers operating under either standard are essentially timed in the same manner, the terminology will have some differences. Like all technology, controllers continue to evolve, and there are now many hybrid versions of the NEMA and Type 170 standards:

- ***NEMA TS-1 Controller*** is an older style that is connected to TS-1 cabinet devices through three MS-type connectors (designated A, B, and C) with a designated pin configuration. A D-connector is typically present to provide more features, but it is not NEMA specified. (Each supplier has developed a different D-connector.) Most modern controllers are TS-2, but some are operated (in a hybrid mode) in TS-1 cabinets and do not take advantage of all TS-2 features.

- ***NEMA TS-2 Type 1 Controller*** is the current NEMA controller type, leveraging all of the capabilities and features of the TS-2 standard. This controller only works in a TS-2 cabinet with synchronous data link control communication.
- ***NEMA TS-2 Type 2 Controller*** is compatible with both TS-2 and TS-1 style cabinets.
- ***Type 170 Controller*** is an older/original style with standardized hardware. It uses Caltrans series cabinets (i.e., 332, 334, and 336) through a C1 connector.
- ***Type 2070 Controller*** is evolved from the 170 standard, but will work in Caltrans cabinets or any NEMA style cabinet. (Note that adaptors and/or interface cards are required, as well as a controller firmware appropriate to the configuration.)
- ***Advanced Transportation Controller (ATC)*** represents the next generation of controller. The standard provides an open-architecture software platform that acts as a universal interface between application programs and the ATC controller units.

The signal designer should ensure that all inputs and outputs can be accommodated by the controller (and style of cabinet). Modified cabinets that allow for additional hardware and/or other intelligent transportation system (ITS) devices might need to be procured.

4.2.3 Detector Cards

Adequate detector rack space should be provided to allow for near-term and possible long-term needs.

Detector cards (which are also referred to as "detector amplifiers" when used with inductive loops) identify user actuations from the field detectors and pass the information along to the signal controller. Most detector cards can handle between one and four detector channels and various modes of operation. Connections and capabilities of detector cards will vary with the method of detection technology being used, but typically work with standard detector racks/input files, provided sufficient slots are available in the signal cabinet.

4.2.4 Flasher and Flash Transfer Relays

The flasher makes it possible for the displays to flash 50 to 60 times per minute (i.e., 50 percent time power-on, plus or minus 5 percent). This flashing power is used for flashing don't walk, flashing yellow arrows, night-time flash, and other flashing indications as needed. The flash transfer relays are used by the signal monitor to disconnect the controller and transfer the signal to emergency flashing operations. Emergency flash operations (typically all indications flashing red) must be configured through appropriate cabinet wiring.

4.2.5 DC Power

Most cabinet equipment (e.g., detector card racks, communications equipment, and other auxiliary equipment) operates in a 12/24-volt DC environment. DC power equipment is handled in a variety of manners depending on the type of cabinet and traffic signal controller.

4.2.6 Load Switches

Load switches are essentially electrical gating (or relay) devices that allow the controller, which operates in a 12/24-volt DC environment, to direct a 120-volt AC current to various signal displays. (Note that load switches are called "switch packs" in a Caltrans series installation.) Each load switch (and associated wiring) plugs into load switch bays in the back panel of the cabinet. The number of load switch bays will dictate the number of output channels that the signal designer has available at the intersection. A load switch is typically required for each vehicle signal phase, each pedestrian phase, and each overlap.

Ensuring that there are enough load switch bays for existing and future phasing is recommended.

4.2.7 Signal Monitor

Modern traffic signal monitors have many advanced features to improve safety and enhance maintenance of traffic signals. The NEMA malfunction management unit (MMU) and the older conflict monitoring unit (CMU) are important safety devices. While these devices are often called conflict monitors, they have evolved from simple conflict and voltage monitoring to enhanced fault monitoring. Enhanced monitors have features beyond monitoring conflicting phases or conflicting indications in a signal head. The enhanced monitors detect and respond to improper voltages caused by malfunctions of the controller unit (CU), load switches, or incorrect wiring of the cabinet. They will identify the type of fault (e.g., conflict, red fail, clearance fail, dual indication, or communications faults in the cabinet) and which signal heads were active at the time of the fault, and they can retrieve historical data about the fault. The monitors will remain in fault mode until reset.

4.2.8 Uninterruptible Power Supply Backup

There are some locations where an uninterruptible power supply (UPS)/battery backup system (BBS) may be necessary or desirable when utility power is not available. A UPS can provide emergency power to connected equipment by supplying power from a separate source (e.g., batteries, solar, or wind). The system can also function as a power conditioner and/or voltage regulation device.

The selection of a UPS should be based on agency objectives and system needs.

Installing UPS systems at locations where there have been power issues helps reduce downtime and electrical damage to equipment. For example, a signalized intersection that is equipped with a UPS can continue to operate through short-term power losses. MUTCD guidance states "except for traffic control signals interconnected with light rail transit systems, traffic control signals with railroad preemption or coordinated with flashing-light signal systems should be provided with a backup power supply" (*1*). UPS systems consist of an enclosure or cabinet, the batteries, the power inverter/conditioner, a battery charger (usually integral to the inverter), and automatic and manual bypass switches. It is also desirable to have an external flashing light to indicate whether the signal is operating on commercial power or UPS (without opening the cabinet).

UPS systems can operate in a variety of modes (i.e., fully operational, dimmed, flashing, or variable), that can affect the length of time a battery charge will last.

4.3 DISPLAYS

Once detection has been processed by the signal cabinet equipment, the signal displays can be used to direct intersection users. The MUTCD contains guidance on traffic signal display types and positions for the most common applications (*1*). For a safe and effective signal design, the displays must support initial and future phasing, as well as be positioned for maximum visibility for each group of users. The displays, support structure, and wiring should reflect consideration of both near-term and long-term needs in order to provide maximum timing flexibility.

4.3.1 Vehicle Displays

Vehicle displays are the most common type of indication at signalized intersections. While most practitioners are familiar with them, the following sections provide guidance on the more atypical displays and situations that may need to be addressed during design.

4.3.1.1 Display Visibility

All displays at an intersection should be easy to see and interpret. This is particularly important when it comes to vehicle displays because there are usually more of them at an intersection than other types of displays. It is important that intersection users establish line-of-sight as they approach an intersection and correctly interpret what they see. Use of strategic signal head placement can ensure line-of-sight is available regardless of the horizontal and vertical geometry on an approach (as shown in Exhibit 4-12 and Exhibit 4-13).

Exhibit 4-12 Advance Signal Placement to Address Horizontal Curve

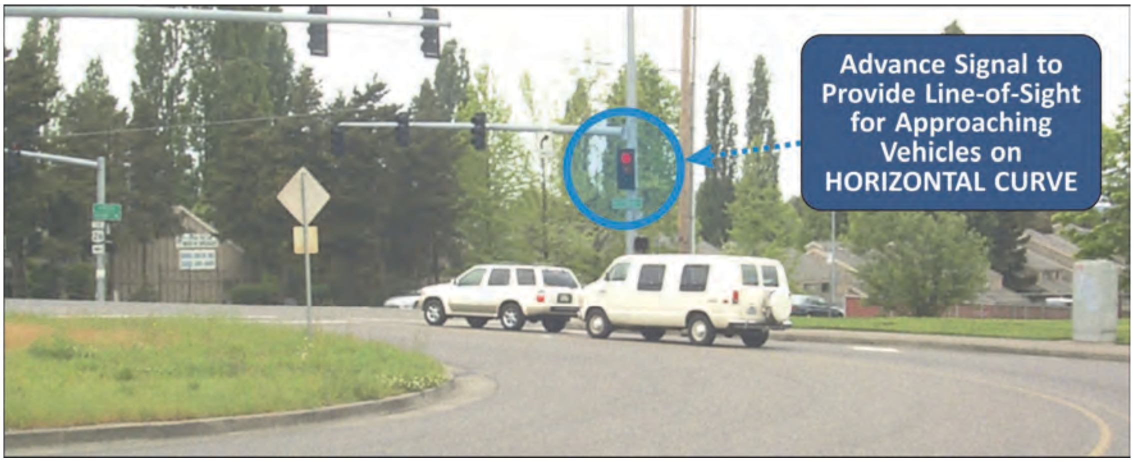


Exhibit 4-13 Advance Signal Placement to Address Vertical Curve

Intersections with limited visibility and/or high-speed approaches may require advance warning "signal ahead" signs. While some beacons flash continuously, they are more effective when they start a few seconds before the onset of yellow. Some controllers have built-in warning systems to determine when to activate the sign prior to terminating the phase, but most systems use a simple overlap to turn on the sign at a predetermined time before the yellow. These timed overlaps occur after a gap in traffic is detected, reducing the effectiveness of decision zone detection because vehicles may arrive during the timed overlap. Some controllers have smart features to overcome this limitation.

Avoiding confusing or contradictory vehicular displays is also important when designing a signal. Programmable signal heads and signal louvers are treatments that can be used to focus the signal indication in the direction of the desired users and away from users for whom the indication is not intended (as depicted in Exhibit 4-14).

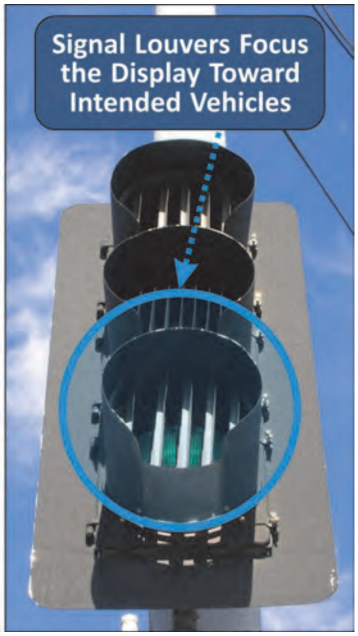


Exhibit 4-14
Programmable Vehicle Signal Heads and Signal Louvers

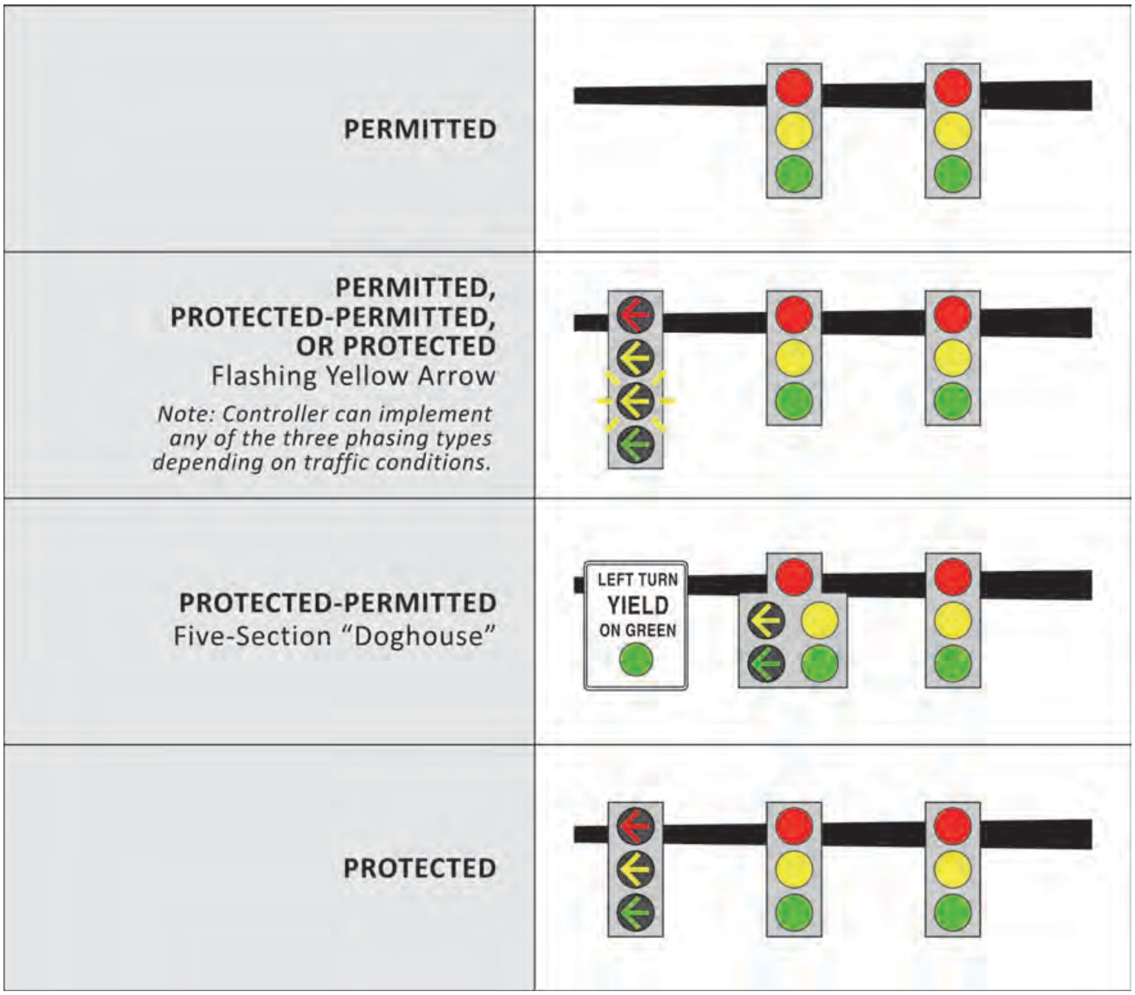
4.3.1.2 Left-Turn Displays

Selecting left-turn signal phasing is the most common phasing decision to be made during design of a signalized intersection. States and local agencies often have criteria for selecting specific signal phasing. Many of these guidelines indicate that a left-turn phase can be justified based on factors that ultimately tie back to the operational or safety benefits derived, including the following:

- Left-turn and opposing through volumes,
- Number of opposing through and turn lanes,
- Cycle length,
- Speed of opposing traffic,
- Sight distance, and
- Crash history.

There are five options for left-turn signal phasing at an intersection (described in detail in Chapter 5): permitted, protected, protected-permitted, split phasing, and prohibited. Exhibit 4-15 depicts various combinations of permitted, protected-permitted, and protected left-turn signal phasing displays. Split phasing uses the same type of displays as protected phasing, but requires additional programming in the controller. Prohibited movements will not require any displays. From a design perspective, wiring and signal cabinet equipment should be provided to support potential signal phasing changes.

Exhibit 4-15 Left-Turn Signal Displays



The flowchart shown in Exhibit 4-16 (in combination with material presented in Exhibit 4-17 and Exhibit 4-18) can help a practitioner determine whether a separate left-turn phase is needed and whether the operational mode should be permitted, protected-permitted, or protected. The flowchart provides a structured evaluation procedure that promotes consistent application of left-turn signal phasing. The guidelines in Exhibit 4-16 were derived from a variety of sources (*10*, *11*, *12*) and require separate evaluation of each left-turn movement on the subject road. The main objective of the flowchart is to identify the least-restrictive left-turn operational mode that can meet desired operational and safety objectives.

Signal Timing Manual, Second Edition

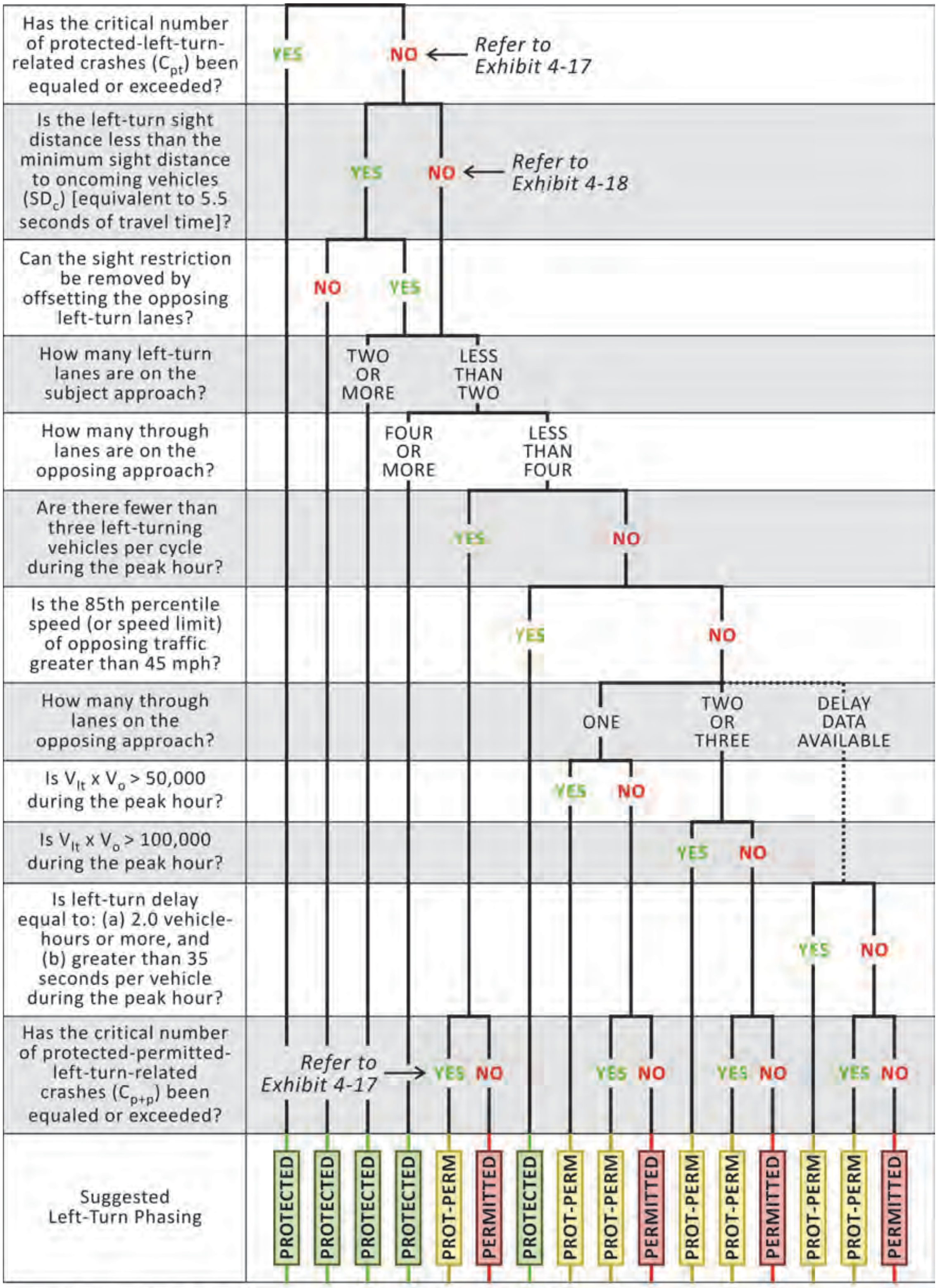


Exhibit 4-16 Left-Turn Phasing Guidelines

V_{lt} = Left-turn volume on subject approach (vehicles per hour)
V_o = Through plus right-turn volume on approach opposing subject left-turn movement (veh per hour)
Prot-Perm = Protected-Permitted (Desirable) or Protected

Source: Adapted from the *Manual of Traffic Signal Design*, 2nd Edition (11), *The Traffic Signal Book* (12), and the *Traffic Engineering Manual* (13).

Exhibit 4-17 Critical Left-Turn-Related Crash Count

Number of Left-Turn Movements on Subject Road	Period during which Crashes Are Considered (Years)	Critical Left-Turn-Related Crash Count (Crashes Per Period)	
		When Considering Protected-Only (C_{pt})	*When Considering Protected-Permitted (C_{p+p})*
One	1	6	4
	2	11	6
	3	14	7
Two	1	11	6
	2	18	9
	3	26	13

Exhibit 4-18 Minimum Sight Distance to Oncoming Vehicles

Oncoming Traffic Speed Limit (Miles Per Hour)	Minimum Sight Distance to Oncoming Vehicles (SD_c) (Feet)
25	200
30	240
35	280
40	320
45	360
50	400
55	440
60	480

In order to account for the inherent variability of crash data, the critical left-turn crash counts identified in Exhibit 4-17 are based on an underlying average critical crash frequency. The underlying averages are 1.3 crashes per year and 3.0 crashes per year for protected-permitted and protected-only left-turn phasing, respectively. If the reported crash count for existing operations exceeds the critical value, then it is likely that the subject intersection has an average left-turn crash frequency that exceeds the aforementioned average (5 percent chance of error), and a more restrictive operational mode would likely improve the safety of the left-turn maneuver.

Note that the flowchart has two alternative paths following the check of opposing traffic speed. One path requires knowledge of left-turn delay; the other requires knowledge of the left-turn and opposing through volumes. The left-turn delay referred to in the flowchart is the delay incurred when no left-turn phase is provided (i.e., the left-turn movement operates in the permitted mode).

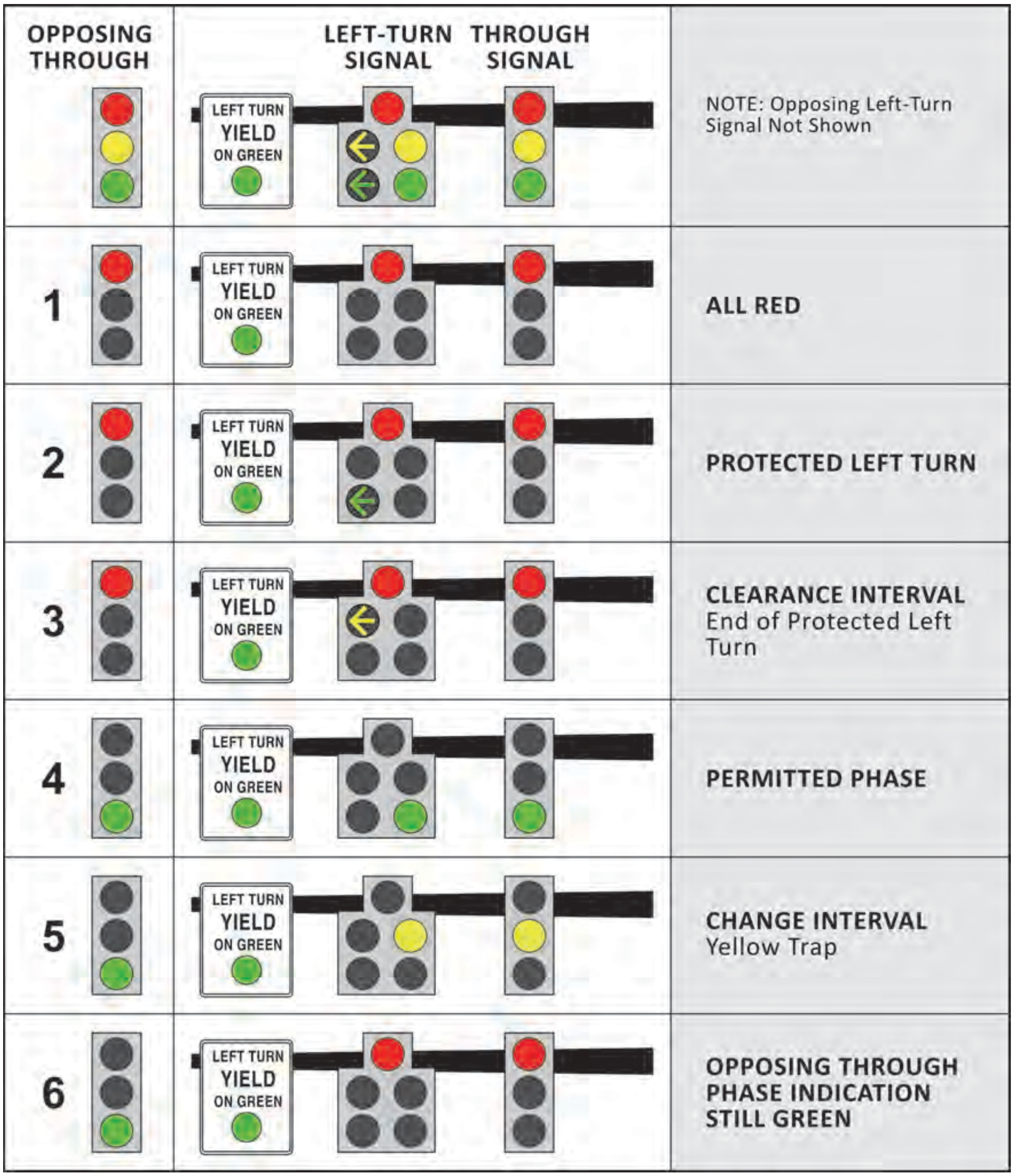
4.3.1.3 Flashing Yellow Arrow Displays

One protected-permitted left-turn display that warrants additional discussion is the recently introduced flashing yellow arrow (FYA) display. (Refer to Chapter 5 for additional guidance on protected-permitted operations.) This indication features a flashing yellow output, which must be accommodated in the signal design with

- Individual wires to/from the cabinet to each of the four indications in the FYA signal head;
- A signal monitor with the functionality to accommodate the FYA indication; and
- An additional load switch for the flashing yellow indication or an unused load switch channel on a pedestrian load switch (the unused yellow output), as load switches only contain three outputs.

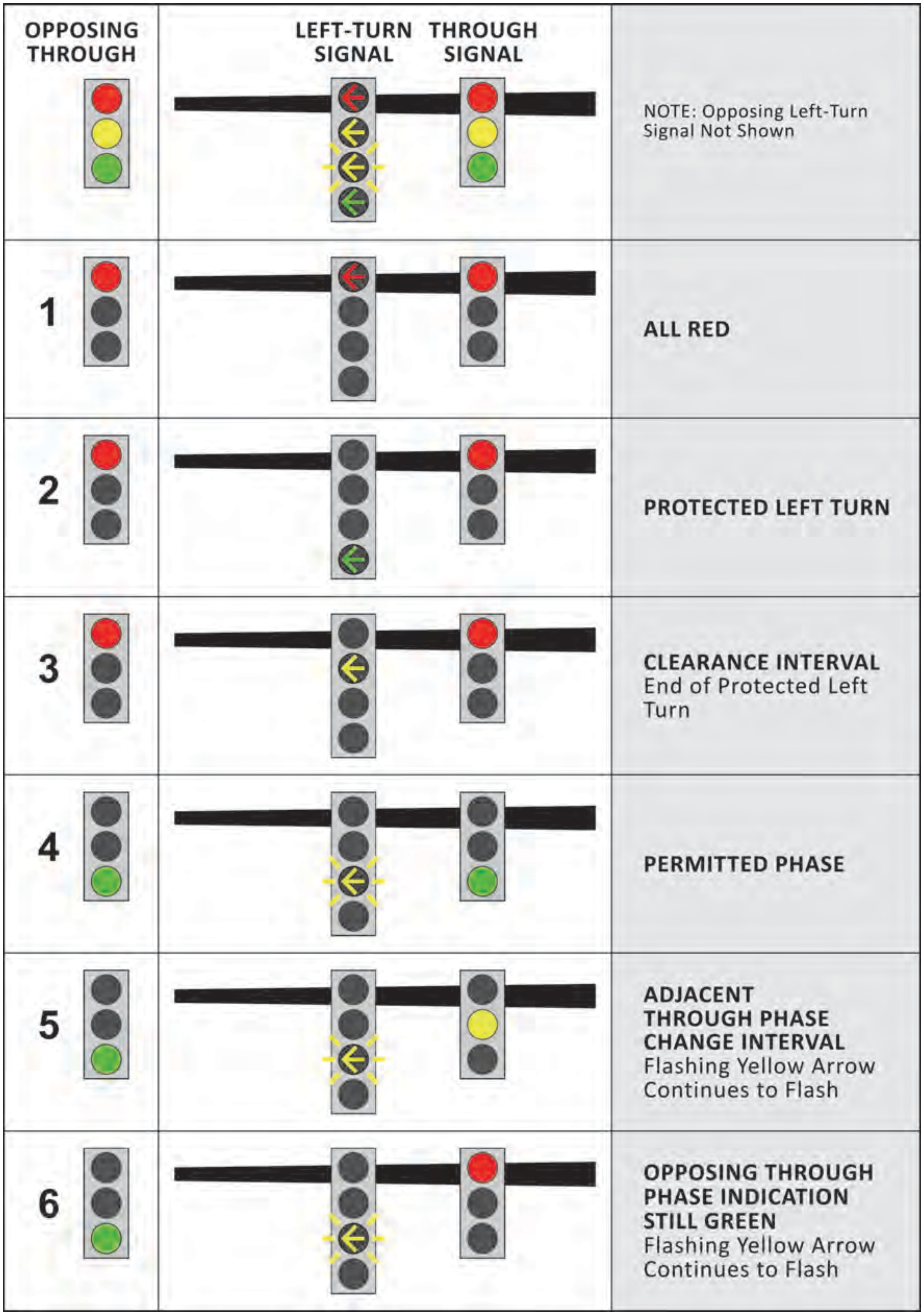
Practice has shown that the FYA indication is more effective than other permitted indications (i.e., steady green ball) at reducing the critical false "go" interpretation by users and resulting yellow trap conditions (*14*, *15*). The yellow trap is a condition where a left-turning user interprets the onset of a steady yellow ball indication and incorrectly assumes oncoming through traffic sees the same steady yellow ball indication. This can be problematic if the left-turning user attempts to "sneak" through the intersection on yellow when oncoming traffic still sees a green indication. Exhibit 4-19 illustrates the yellow trap that can occur with a doghouse signal, and Exhibit 4-20 illustrates how an FYA can mitigate that condition.

Exhibit 4-19
Illustration of the Yellow Trap



Source: Adapted from *Signalized Intersections: An Informational Guide* (*16*)

Exhibit 4-20
Illustration of the Flashing Yellow Arrow



There are signal timing mitigations other than the FYA that can be implemented to help with the yellow trap problem. For example, the yellow trap could be mitigated

Signal Timing Manual, Second Edition

through the installation of signs, forcing a minor street call, or installing a flashing red arrow (FRA). The FRA indication is an alternative to the FYA (that needs controller firmware and a signal monitor that is compatible with the FRA), which requires vehicles to make a complete stop prior to making their movement.

4.3.1.4 Right-Turn Displays

Permitted, protected-permitted, and protected displays are all options to support various right-turn signal phasing. The simplest display (permitted) should be used unless more complex designs are necessary to improve capacity or clarify complex phasing. If protected phasing is necessary (or may be necessary in the future), signal designers should provide a separate overlap load switch (or load bay position for a future load switch) for right-turn displays used for exclusive right-turn lanes, rather than tying the right-turn arrow indication to the compatible left-turn signal phase. The latter is often done to simplify wiring and reduce load switch channels, but sacrifices flexibility in signal timing, which may result in less effective traffic operations. Pedestrian conflicts and right-turn-on-red laws should also be considered when selecting the type of signal indication to use for a right-turn movement.

4.3.2 Pedestrian Displays

There are two typical styles of pedestrian displays (shown in Exhibit 4-21): (1) those with a countdown timer for the flashing don't walk (FDW) interval and (2) those without any countdown timer. Signal timing requirements are equivalent for both pedestrian signal displays. The reader should refer to the MUTCD for determining the use of countdown pedestrian displays (*1*).

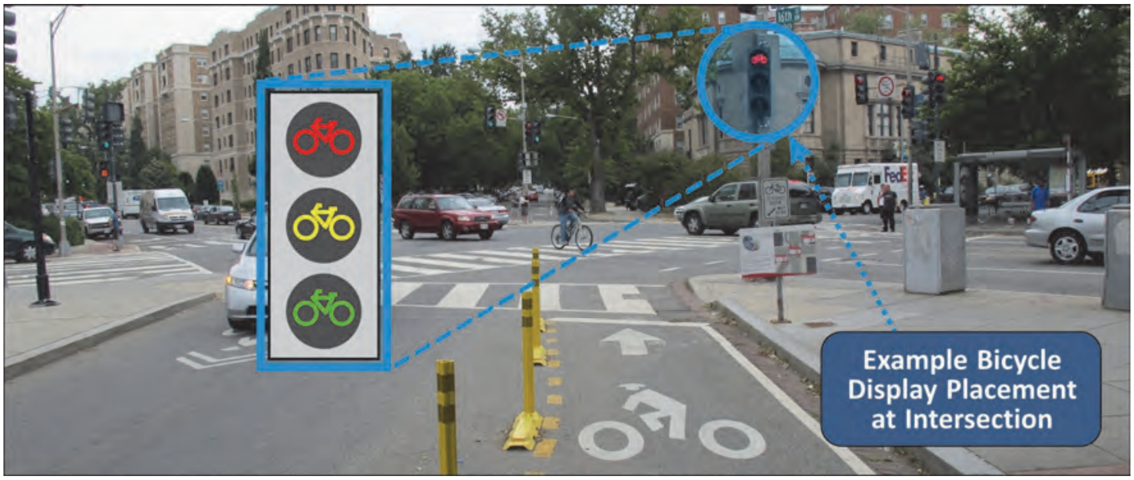
Exhibit 4-21 Typical Pedestrian Signal Displays

Source: Adapted from the MUTCD (*1*)

4.3.3 Bicycle Displays

The signal designer should minimize conflicts for bicycles while retaining effective operations. Bicycle movements can be served concurrently with motorized vehicle phases (and/or pedestrian phases) or served independently with a separate bicycle signal phase. Experimental use of bicycle displays (following FHWA guidance) is in place, but the displays are not part of the 2009 MUTCD. Bicycle displays tend to be used in locations where conveying special signal phasing for bicycle movements (and/or clarification of traffic control for bicycle movements) is advantageous to the desired outcomes of the signalized intersection operations. Exhibit 4-22 shows an example of a bicycle display, per the California MUTCD (*17*), and an example of the placement of a near-side bicycle display at an intersection.

Exhibit 4-22 Example Bicycle Display (Caltrans) and Placement at Intersection



4.3.4 Transit Displays

If exclusive transit phasing is needed (or is likely to be needed in the future), signal cabinet and controller selection should explicitly consider this need, as well as conduit sizing. Exhibit 4-23 shows alternative LRT displays, which can also be used for exclusive transit movements. Note that different jurisdictions will have different standards for transit displays, so agency standards should be verified before installation.

Exhibit 4-23 Typical Light Rail Signal Displays for a Single LRT Route

THREE-LENS LIGHT RAIL SIGNAL	STOP	
	PREPARE TO STOP	Flashing Lens
	GO	
TWO-LENS LIGHT RAIL SIGNAL	STOP	
	GO	Lens May Be Used in Flashing Mode to Indicate Prepare to Stop

Source: Adapted from the MUTCD (*1*)

4.4 SIGNALIZED SYSTEM DESIGN

Once individual intersections have been designed with detection, cabinets, and displays, they often need a way to transmit information and work as a system. From a maintenance and operations perspective, every intersection should have some form of communication because it is necessary for updating the local clock, monitoring faults, accessing performance data, and remotely making timing adjustments. Communication can be achieved by interconnecting the signal controllers to a field master/master controller or connecting the signal controllers to a central system that directly monitors them. A system of traffic signals is typically composed of the following (illustrated in Exhibit 4-24):

- A series of local controllers,
- Communications hardware (e.g., conduit and wiring) and software, and
- A "master" controller or "central" system.

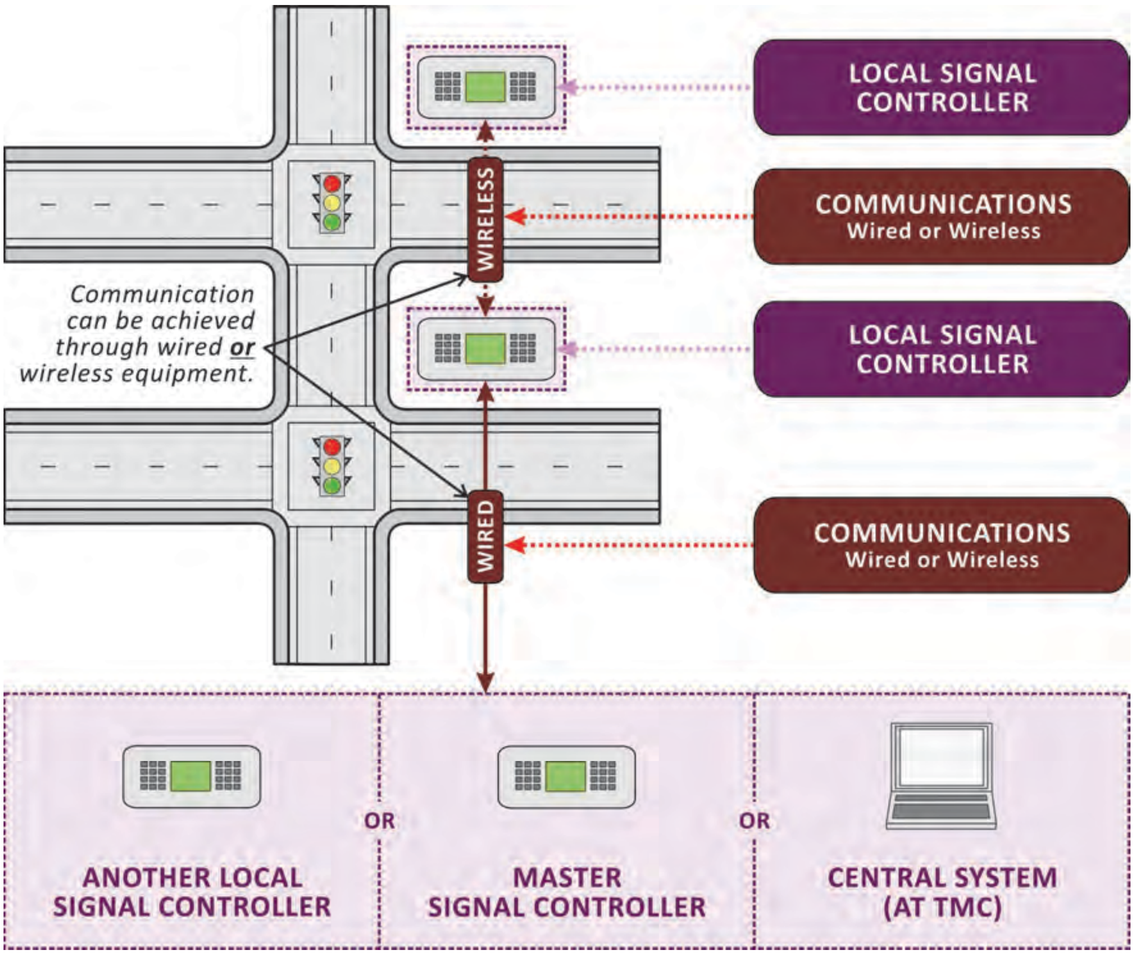


Exhibit 4-24 Physical Components of an Interconnected Signal System

A communications system allows a practitioner to monitor signals from a remote location and upload/download signal timing information from each individual intersection controller; it also allows the system to sync for time-of-day based control. Streamed video images can significantly change the communications requirements of a system, but provide the ability for practitioners to visually monitor signal operations and timing remotely.

There are many alternative system configurations that may be appropriate based on the size of the system and the nature of the communications. If a group of signals is using closed-loop operations, the individual intersection controllers do not communicate with a central system; they communicate with a master controller (which can be configured to communicate with a central system). Conversely, a system can be set up so that all intersections communicate with a central system, which is generally located at a traffic management center (TMC) or signal shop.

Signal timing plans will typically reside in the local controllers in the field, but may also reside in a central system database. Timing plans for most systems that do not directly control intersections from a central system can have timing parameters adjusted by (1) hand-keying new parameters directly into the controller or (2) downloading new elements from PC-based software (e.g., through a direct serial connection, Ethernet connection, or dial-up modem). More information about uploading and downloading timing plans is available in Chapter 8.

4.5 COMPREHENSIVE DESIGN CONSIDERATIONS

The traffic signal design and layout should support as much signal timing flexibility as possible for existing and future conditions.

If possible, a practitioner should always design a signal with future conditions in mind. Although signal designs should be consistent with the agency objectives and system needs, a comprehensive signal design would

- Provide detection for all movements and modes—including motorized vehicles, non-motorized vehicles, pedestrians, and transit users—which allows for (1) fully-actuated or "free" signal operations and (2) measuring the performance for all movements and modes.
- Provide advance detection for higher speed approaches to allow for decision zone protection or more efficient green extension.
- Avoid lane use that requires specific signal phasing, such as shared left-through lanes (typically requiring split phasing), where possible.
- Incorporate mast arm lengths and signal poles to support future signal displays (e.g., typically extend mast arms to the center of the farthest left-turn lane even if a left-turn display is not required upon construction).
- Include communications between the traffic management location (center and/or shop) to the controllers for monitoring of signal operations and adjustment of signal timing.
- Provide for additional capacity in the terminal facilities, detector slots, underground conduits, and other cabinet space necessary to support future signal timing, communications, and automated data collection/performance measurement needs.

Some lessons learned, specific to future needs and possible equipment problems, include the following:

- Providing enough conduit space (e.g., two 3-inch conduits with ducts across approaches from the cabinet and one 3-inch conduit across other approaches).
- Locating pull boxes and pole foundations outside of possible future widening.
- Installing spare conductors for unused load bay positions for future load switches, plus at least 10 percent spares for conductors that fail.

If one of the critical paths fails, having excess capacity available will allow a practitioner to get a signal back in service more quickly.

4.6 REFERENCES

1. *Manual on Uniform Traffic Control Devices for Streets and Highways,* 2009 Edition. United States Department of Transportation, Federal Highway Administration, Washington, D.C., 2009.
2. Klein, L. A., M. K. Mills, and D. R. P. Gibson. *Traffic Detector Handbook,* 3rd Edition. Report FHWA-HRT-06-108, Federal Highway Administration, United States Department of Transportation, 2006.
3. McGee, Sr., H., K. Moriarty, K. Eccles, M. Liu, T. Gates, and R. Retting. *NCHRP Report 731: Guidelines for Timing Yellow and All-Red Intervals at Signalized Intersections*. Transportation Research Board of the National Academies, Washington, D.C., 2012.
4. Parsonson, P. S., R. W. Roseveare, and J. R. Thomas, Jr. Small-Area Detection at Intersection Approaches. *Traffic Engineering*, Vol. 44, No. 5, 1974, pp. 8–17.
5. Zegeer, C. V., and R. C. Deen. Green Extension Systems at High-Speed Intersections. *ITE Journal*, Vol. 48, No. 11, 1978, pp. 19–24.
6. Chang, M. S., C. J. Messer, and A. J. Santiago. Timing Traffic Signal Change Intervals Based on Driver Behavior. In *Transportation Research Record 1027*, TRB, National Research Council, Washington, D.C., 1985, pp. 20–30.
7. *A Policy on Geometric Design of Highways and Streets,* 6th Edition. American Association of State Highway and Transportation Officials, Washington, D.C., 2011.
8. Bonneson, J. A., and P. T. McCoy. *Manual of Traffic Detector Design,* 1st Edition. Institute of Transportation Engineers, Washington, D.C., 1994.
9. Lin, F. B. Optimal Timing Settings and Detector Lengths of Presence Mode Full-Actuated Control. In *Transportation Research Record 1010*, TRB, National Research Council, Washington, D.C., 1985, pp. 37–45.
10. Zegeer, C. V., C. Seiderman, P. Lagerwey, M. Cynecki, M. Ronkin, and R. Schneider. *Pedestrian Facilities User Guide—Providing Safety and Mobility*. Report FHWA-RD-01-102, Federal Highway Administration, United States Department of Transportation, 2002.
11. Kell, J. H., and I. J. Fullerton. *Manual of Traffic Signal Design*. Institute of Transportation Engineers, Washington, D.C., 1982.
12. Orcutt, Jr., F. L. *The Traffic Signal Book*. Prentice Hall, Englewood Cliffs, New Jersey, 1993.
13. Chapter 3: Signals. In *Traffic Engineering Manual*. Florida Department of Transportation, Tallahassee, Florida, 1999, pp. 121–159.
14. Brehmer, C. L., K. C. Kacir, D. A. Noyce, and M. P. Manser. *NCHRP Report 493: Evaluation of Traffic Signal Displays for Protected/Permissive Left-Turn Control*. Transportation Research Board of the National Academies, Washington, D.C., 2003.
15. Noyce, D. A., C. R. Bergh, and J. R. Chapman. *NCHRP Web-Only Document 123: Evaluation of the Flashing Yellow Arrow Permissive-Only Left-Turn Indication Field Implementation*. Transportation Research Board of the National Academies, Washington, D.C., 2007.

16. Chandler, B. E., M. C. Myers, J. E. Atkinson, T. E. Bryer, R. Retting, J. Smithline, J. T. P. Wojtkiewicz, G. B. Thomas, S. P. Venglar, S. Sunkari, B. J. Malone, and P. Izadpanah. *Signalized Intersections: An Informational Guide*, Second Edition. Report FHWA-SA-13-027, Federal Highway Administration, United States Department of Transportation, 2013.
17. *California Manual on Uniform Traffic Control Devices,* 2012 Edition. California Department of Transportation, 2012.

CHAPTER 5
INTRODUCTION TO TIMING PLANS

CONTENTS

LIST OF EXHIBITS

CHAPTER 5. INTRODUCTION TO TIMING PLANS

Developing signal timing plans involves selecting timing values that ultimately determine how an intersection (or more commonly a system of intersections) operates.

Chapters 5, 6, and 7 make up a three-part series about developing signal timing plans. Chapter 5 describes basic signal timing concepts that a practitioner should understand before defining signal timing values. Chapter 6 provides detailed information about signal timing parameters required at every signalized intersection, and Chapter 7 describes the timing parameters that must be defined when signalized intersections are coordinated. Using the information from Chapters 5, 6, and 7, a practitioner should be able to develop a timing plan that meets a defined timing strategy and local operational objectives (discussed in Chapter 3).

5.1 BASIC SIGNAL TIMING CONCEPTS

There are many signal timing parameters that must be defined for every user group at an intersection for every time period throughout the day. In order to keep the parameters organized, practitioners have developed conventions for how movements are referenced, how phases are numbered, how overlaps work, and how those movements, phases, and overlaps correspond to detectors, signal cabinet equipment, and displays.

5.1.1 Movement and Phase Numbering

When developing a signal timing plan, movements should be the first focus. Signal phasing can be assigned to carry out the desired operations.

Movements describe user actions at an intersection. At a signalized intersection (with four approaches), it is possible to have twelve one-way vehicular movements and four two-way pedestrian movements. Each of these movements can be assigned a number for reference. The *Highway Capacity Manual* (HCM) (*1*) assigns movement numbers as shown in Exhibit 5-1 (illustrated by the gray squares). The HCM gives each right-turn movement its own number (separate from the through movement) by adding 10 to the adjacent through movement number. It is important for a practitioner to understand the difference between lane assignments and movements. Note that a single movement can be accommodated by multiple lanes (e.g., through movement in two lanes), or multiple movements can be accommodated by a single lane (e.g., through/right-turn lane).

A traffic signal phase is a timing process, within the signal controller, that facilitates serving one or more movements at the same time (for one or more modes of users). A practitioner must assign phase numbers to the movements at a signalized intersection in order to begin selecting signal timing values. A typical four-legged intersection with protected left-turn movements (protected movements have the right-of-way over other movements) will generally follow the phase numbering shown in Exhibit 5-1 (illustrated by the blue boxes). This standard National Electric Manufacturers Association (NEMA) phase numbering system combines the through movements with the right-turn movements, which are typically permitted (meaning they can be made after yielding to conflicting bicycle and pedestrian movements). Occasionally, the right-turn movement may be a protected movement and timed using an overlap, which is discussed in Section 5.1.4.

To further explain the relationship between movements and phases, Exhibit 5-2 illustrates the typical movement and phase numbering used at an intersection with permitted left-turn movements (i.e., no protected left-turn phases). In this scenario, all of the movements on an approach are assigned to one phase. Permitted movements are shown as dashed arrows in Exhibits 5-1 and 5-2. Note that because of the concurrent through vehicle movements and parallel pedestrian phases, right-turn movements in both exhibits are shown as permitted because they must yield to through bicycles and pedestrians.

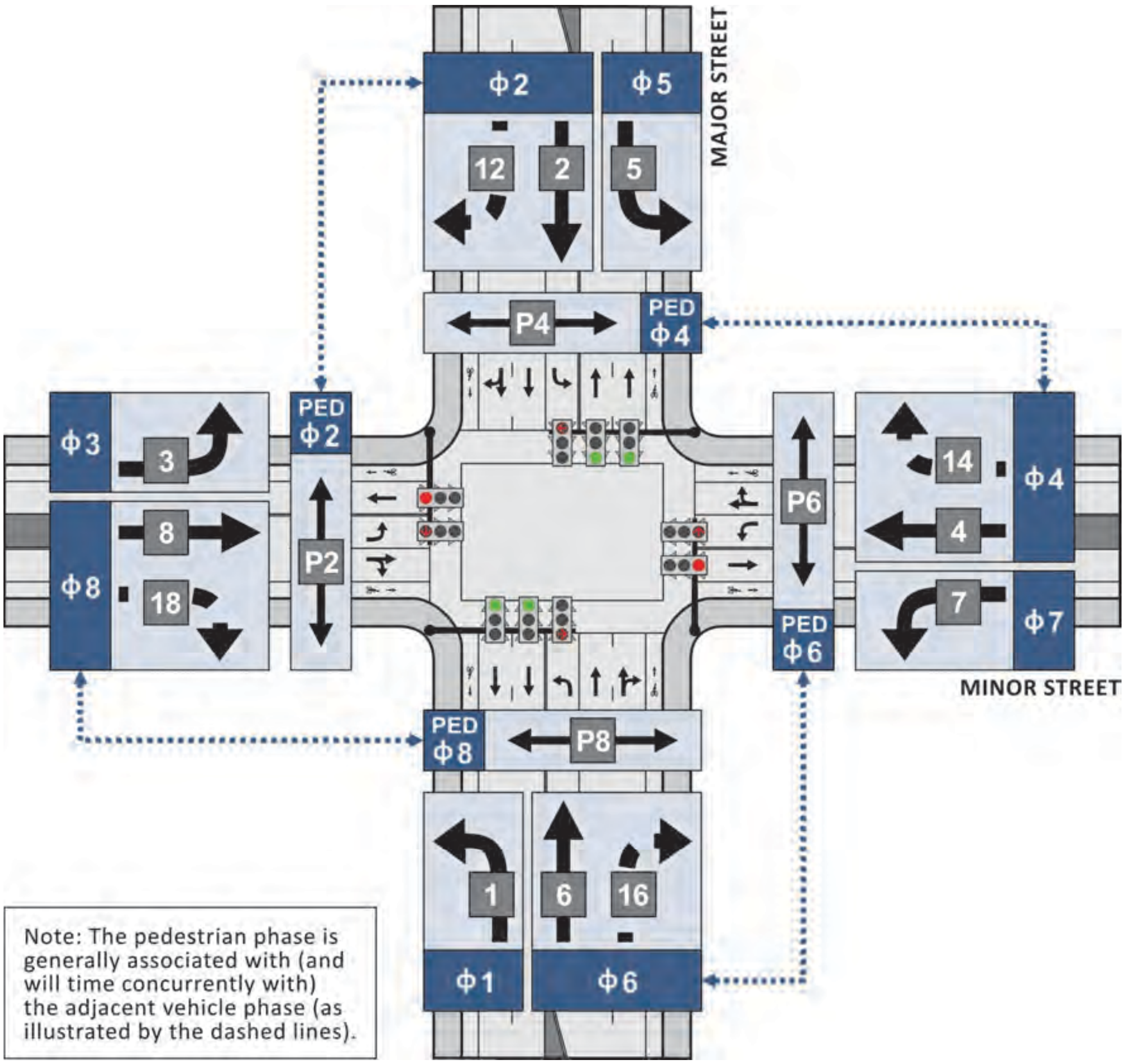


Exhibit 5-1 Typical Movement and Phase Numbering (with Protected Left Turns)

Under a typical phase numbering scheme, there are several conventions that a practitioner should attempt to follow:

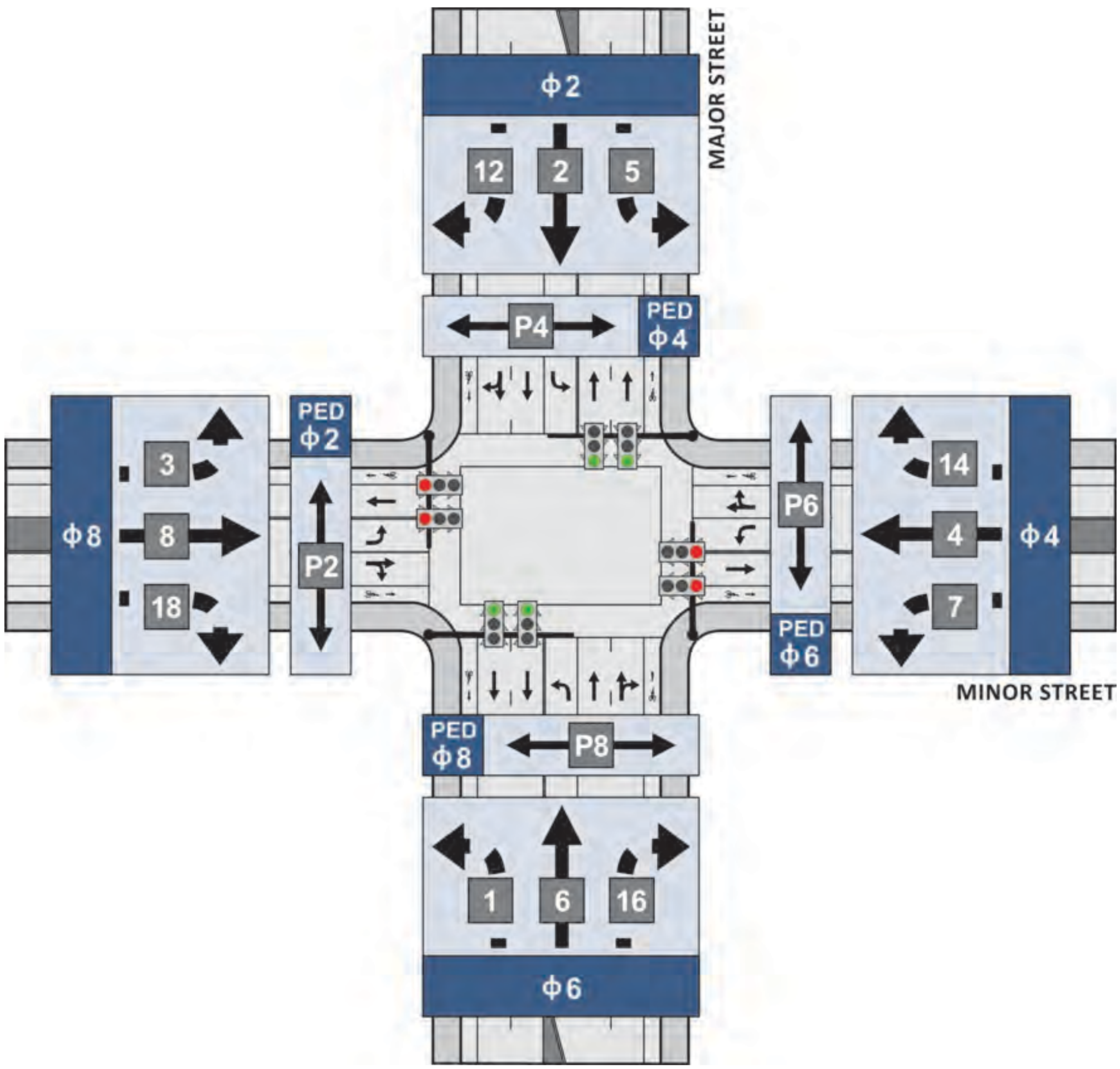
- ***Even*** phases are typically associated with ***through*** movements.
  - Phases 2 and 6 generally represent the major street through movements.
  - Phases 4 and 8 generally represent the minor street through movements.
- ***Odd*** phases are typically associated with ***left-turn*** movements.
  - Phases 1 and 5 generally represent the major street left-turn movements.

- Phases 3 and 7 generally represent the minor street left-turn movements.

- ***Pedestrian*** phases are typically set up to run concurrently with the even-numbered vehicular phases. They are generally assigned the ***same phase number as the adjacent, parallel vehicular phases***.
  - Pedestrian Phases 2 and 6 generally represent the major street pedestrian movements.
  - Pedestrian Phases 4 and 8 generally represent the minor street pedestrian movements.

To avoid confusion, it is common practice to maintain a consistent phase numbering scheme within a specific jurisdiction. For example, Phase 2 could always be defined as the major street through movement in the northbound (or eastbound) direction, or alternatively, as the coordinated phase (regardless of direction). This manual will always follow the convention that Phases 2 and 6 are the major street phases and that they are the coordinated phases when under coordinated operations.

Exhibit 5-2 Typical Movement and Numbering (with Permitted Left Turns)



5.1.2 Ring-and-Barrier Concept

Rings and barriers allow phases to time independently and flexibly, yet within a structure.

Rings and barriers fundamentally define how a controller organizes phases, so that compatible phases can time together and conflicts do not occur.

- ***Ring:*** A ring shows a sequence of conflicting phases. Dual (or two) ring operations allow compatible phases to operate concurrently with (i.e., at the same time as) phases in another ring.
- ***Barrier:*** A barrier is the point at which the phases in both rings must end simultaneously. Barriers typically separate major and minor street phases.

The example ring-and-barrier diagram in Exhibit 5-3 has two rings and two barriers, which organize eight phases. Note that the left-turn phases are protected and leading (i.e., preceding the through movements), which is the most common phasing for an intersection with four approaches and protected left-turn phasing. Other left-turn phasing alternatives are discussed in the following section.

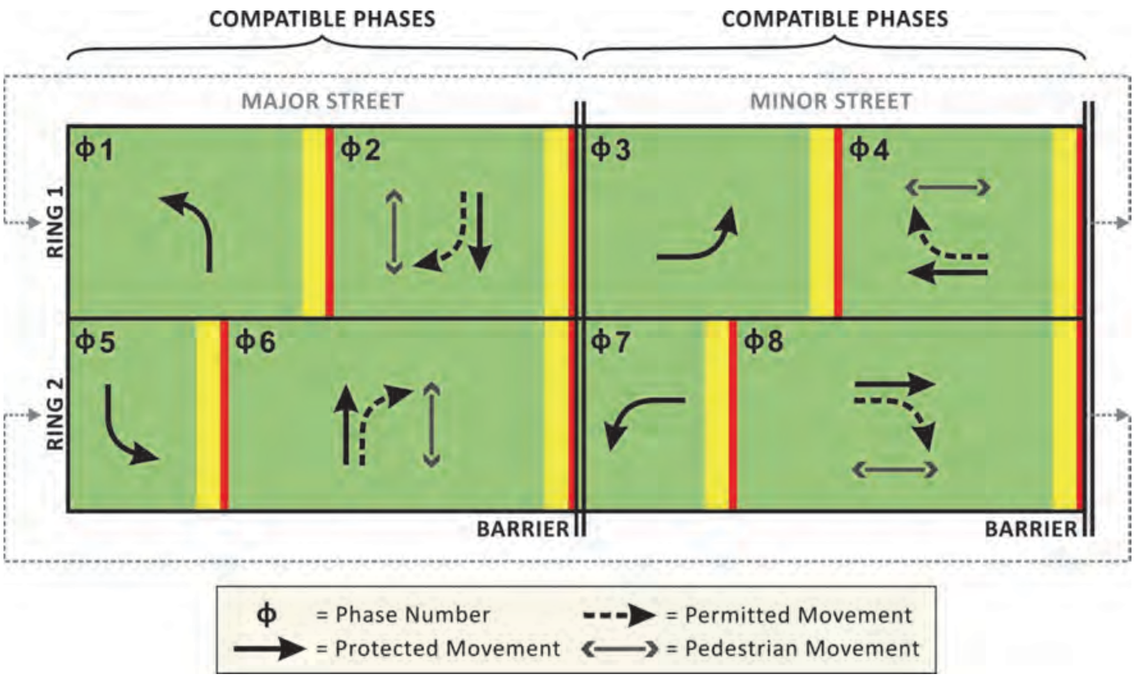


Exhibit 5-3 Basic Ring-and-Barrier Diagram

Phases within two barriers are known as a ***compatible phases*** (as shown in Exhibit 5-3). Between the barriers, several rules apply:

- Any phase in Ring 1 can time with any phase in Ring 2. For example (in Exhibit 5-3), Phases 1, 2, 5, and 6 are compatible phases, so Phases 1 or 2 can time with Phases 5 or 6. Similarly, Phases 3, 4, 7, and 8 are compatible phases, so Phases 3 or 4 can time with Phases 7 or 8.
- Any phase in a ring can be skipped and/or give unused time to a following phase in that ring. For example (in Exhibit 5-3), Phase 1 can give time to Phase 2, and Phase 5 can give time to Phase 6. Alternatively, Phases 3 and 7 can be skipped if there is no demand.
- Subject to additional rules described in Chapter 6, it is possible for only one ring to have a phase timing (i.e., in the other ring, all phases are resting in red).

5.1.3 Left-Turn Phasing

There are five options for left-turn phasing at an intersection (summarized in Exhibit 5-4): permitted, protected, protected-permitted, split phasing, and prohibited. In general, the type of phasing used for one left-turn movement is also used for its opposing left-turn movement. For example, if one left-turn movement is permitted, the opposing left turn is also generally permitted. However, this is not a requirement; the left-turn phasing should be movement-specific and chosen based on a variety of operational and safety factors (discussed in detail in Chapter 4). A practitioner should consult local jurisdiction guidelines when determining which type of left-turn phasing to use at an intersection. Additional information about each left-turn phasing option is provided in the following sections, followed by information on left-turn phase ***sequence*** options.

Exhibit 5-4 Left-Turn Phasing Options

Left-Turn Phasing Option	Description	Advantages	Challenges
Permitted Left-Turn Phase	Served with the adjacent through movement, requiring left-turning vehicles to yield to conflicting vehicle and pedestrian movements	◻ Reduced intersection delay ◻ Efficient green allocation	◻ Requires users to choose acceptable gaps in traffic ◻ Yellow trap can occur if opposing movement is a lagging left turn
Protected Left-Turn Phase	Left-turning vehicles are given the right-of-way without any conflicting movements	◻ Reduced delay for left-turning vehicles ◻ Users always receive exclusive right-of-way; gaps in traffic do not need to be identified	◻ Increased intersection delay
Protected-Permitted Left-Turn Phase	Combination of permitted and protected left-turn phasing; users receive a protected interval, but can also make permitted movements as the conflicting through phase receives a green indication	◻ Compromise between safety of protected left-turn phase and efficiency of permitted left-turn phase ◻ No significant increase in delay for other movements	◻ Fewer options for maximizing progression of through vehicles during coordination (unless flashing yellow arrow displays are used) ◻ Yellow trap can occur if opposing movement is a lagging left turn
Split Phase	Assignment of right-of-way to all movements of a particular approach, followed by all of the movements of the opposing approach	◻ Accommodates use of shared lanes (e.g., left/through lane) ◻ Necessary when opposing left-turn paths overlap because of intersection geometry ◻ Avoids conflict between opposing left-turning vehicles	◻ Increased coordinated cycle length, particularly if both split phases have concurrent pedestrian phases ◻ Less efficient than other types of left-turn phasing
Prohibited Left-Turn Phase	Implemented to maintain mobility at an intersection through use of a “no left turn” sign (particularly during times of day when gaps are unavailable)	◻ Reduced conflicts at the intersection	◻ Users must find alternative routes

5.1.3.1 Permitted Left-Turn Phasing

Permitted left-turn phasing is depicted in the ring-and-barrier diagram in Exhibit 5-5 and has the following characteristics:

- *Right-of-Way:* Permitted phasing requires a user to yield to conflicting vehicular and pedestrian traffic before completing a left turn.
- *Displays:* Both the left-turn and opposing through movements are presented with a circular green indication (i.e., green arrow is never provided).
- *Intersection Conditions:* Permitted operations are primarily used when traffic is light to moderate and sight distance is adequate.
- *Advantages:* This display option provides the most efficient green time allocation, but the efficiency is dependent on the availability of gaps in the conflicting traffic.
- *Challenges:* This mode can have an adverse effect on safety in some situations, such as when the left-turning vehicle's view of conflicting traffic is restricted or when adequate gaps in traffic are not present. The yellow trap (see Chapter 4) can occur if the opposite direction has a lagging left-turn movement.

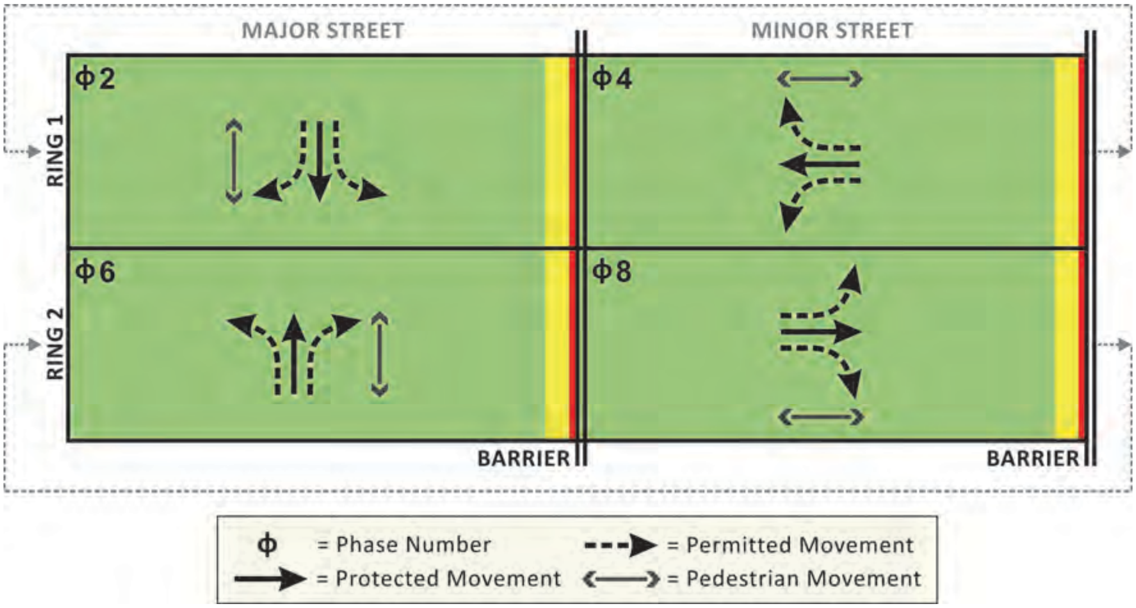


Exhibit 5-5 Ring-and-Barrier Diagram Showing Permitted Left-Turn Phasing

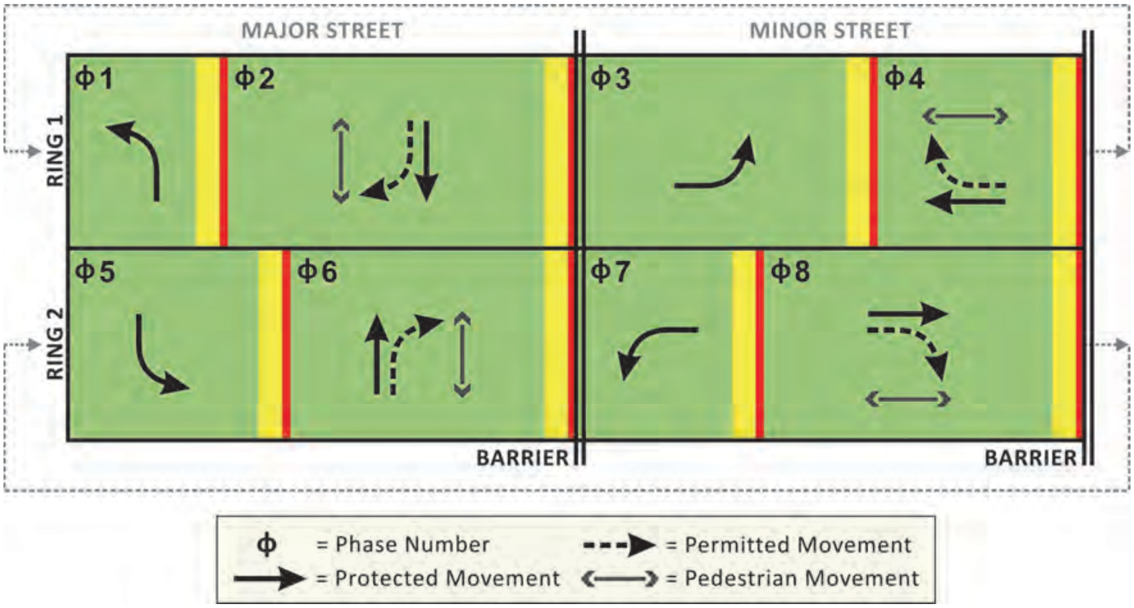
5.1.3.2 Protected Left-Turn Phasing

Protected left-turn operations are shown in Exhibit 5-6 and have the following characteristics:

- *Right-of-Way:* Protected left-turn phasing assigns the right-of-way to users turning left at the intersection.
- *Displays:* It allows turns to be made only on a green arrow display.
- *Intersection Conditions:* An exclusive left-turn lane is typically provided with this phasing.
- *Advantages:* This operation provides for efficient left-turn service and is recognized as the safest type of left-turn operation.

- *Challenges:* The added left-turn phase increases the lost time within the cycle length and may increase delay for other movements.

Exhibit 5-6 Ring-and-Barrier Diagram Showing Protected Left-Turn Phasing



5.1.3.3 Protected-Permitted Left-Turn Phasing

Protected-permitted left-turn operations are shown in Exhibit 5-7 and have the following characteristics:

- *Right-of-Way:* Protected-permitted operations are a combination of the permitted and protected modes. Left-turning vehicles have the right-of-way during the protected left-turn phase and can also complete the left turn "permissively" when the adjacent through movement receives its circular green indication.
- *Displays:* This type of operation requires the use of either a flashing yellow arrow (FYA) display or a five-section “doghouse” display (see Chapter 4 for details).
- *Advantages:* This mode provides for efficient left-turn service, often without causing a significant increase in delay to other movements. This mode also tends to provide a relatively safe left-turn operation as long as adequate sight distance is available.
- *Challenges:* Protected-permitted left-turn phasing should be applied with caution when a phasing sequence other than lead-lead is used (see Section 5.1.3.6 for information on phase sequence). For protected-permitted phasing, the lead-lead sequence prevents the yellow trap associated with lagging left-turn phasing and five-section heads (explained in Chapter 4). However, under light traffic conditions and in the absence of minor street traffic, left-turn phases may be re-serviced without crossing the barrier first, resulting in de facto lagging left turns and a potential yellow trap. Most modern traffic signal controllers have a feature that provides left-turn backup protection. It typically works by omitting the protected left-turn phase when the adjacent through movement is green, ensuring that the left-turn phase will always be preceded by

a barrier. The name of this feature varies by different traffic signal controllers, but some common names include Five-Section Head Restrictions, Five-Section Logic, Trap Protected Phase, Backup Prevent Phases, and Anti-Backup. Protected-permitted operations can also be challenging when used along a coordinated corridor. Additional opportunities for left-turn movements mean fewer opportunities for through vehicles to progress along a corridor. Chapter 7 provides more information on coordination considerations.

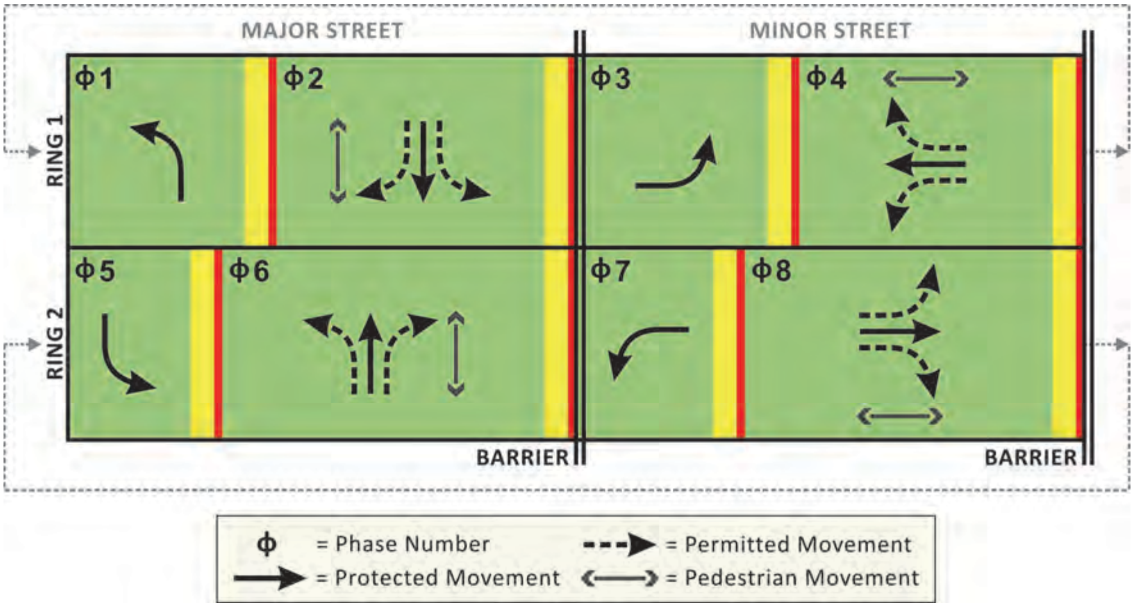


Exhibit 5-7 Ring-and-Barrier Diagram Showing Protected-Permitted Left-Turn Phasing

5.1.3.4 Split Phasing

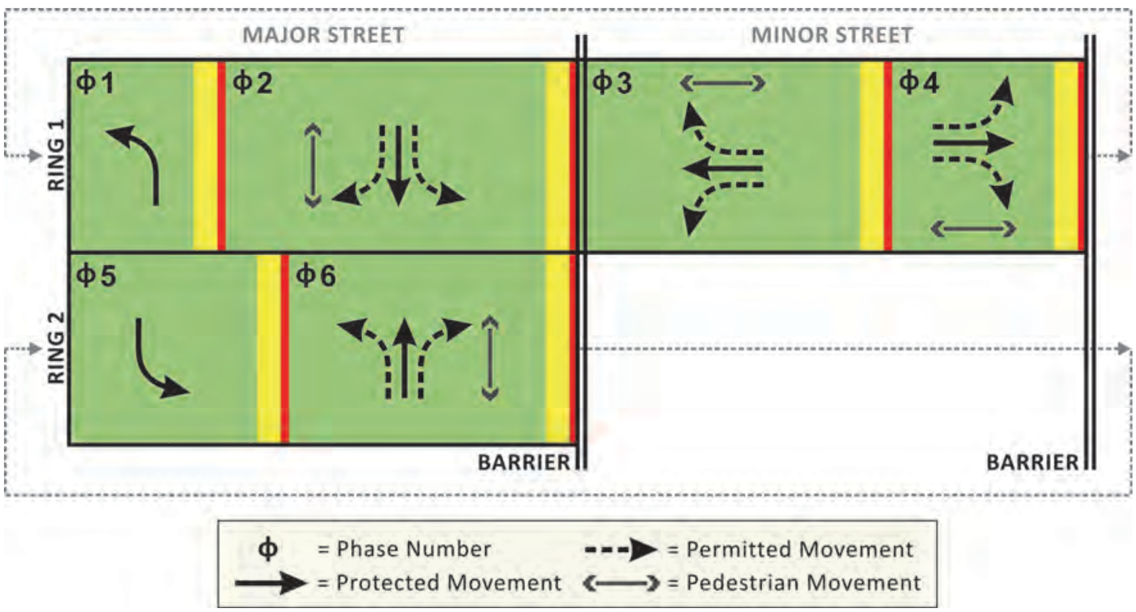
Split phasing is depicted in the ring-and-barrier diagram in Exhibit 5-8 and has the following characteristics:

- *Right-of-Way:* Split phasing assigns right-of-way to all movements on a particular approach, followed by all of the movements on the opposing approach.
- *Displays:* Split phasing uses the same type of displays as protected operations, but requires additional programming in the controller.
- *Intersection Conditions:* Split phasing may be necessary when the following conditions are present (*2*):
    - o There is a need to accommodate one or more left-turn lanes on opposing approaches, but sufficient width is not available to ensure adequate separation between vehicle paths in the middle of the intersection. This condition may also be caused by a large intersection skew angle.
    - o The larger left-turn lane volume is equal to its opposing through lane volume during most hours of the day. ("Lane volume" represents the movement volume divided by the number of lanes serving it.)
    - o The width of the road is constrained such that an approach lane is shared by the left-turn and through movements, yet the left-turn volume is sufficient to justify a left-turn phase.

- One of the two approaches has a heavy volume, the other approach has minimal volume, and actuated control is used. In this situation, the phase associated with the low-volume approach would rarely be called, and the intersection would mostly function as a T-intersection.
- Crash history indicates an unusually large number of sideswipe or head-on crashes in the middle of the intersection involving left-turning vehicles.

- *Advantages:* Split phasing prevents conflicts between opposing left-turning vehicles.
- *Challenges:* This phasing is generally less efficient than other types of left-turn phasing. It typically increases the cycle length or, if the cycle length is fixed, reduces the time available to the intersecting road (*2*).

Exhibit 5-8 Ring-and-Barrier Diagram Showing Split Phasing



5.1.3.5 Prohibited Left-Turn Phasing

Prohibiting left-turn movements is a phasing option with the following characteristics:

- *Right-of-Way:* Left turns can be prohibited during some or all times of day.
- *Displays:* If left turns are prohibited during all times of day, no signal displays are needed, but a “No Left Turn” sign is required. If left turns are prohibited only during certain times of day (when gaps in traffic are unavailable and permitted phasing may be unsafe), displays should be chosen based on the appropriate phasing, and a supplemental sign should be added. Exhibit 5-9 shows an example of such a sign in Toronto, Ontario, where left turns are prohibited during the morning and evening periods.
- *Advantages:* Prohibiting left turns can help to maintain mobility at an intersection.
- *Disadvantages:* Vehicles wanting to make a left turn must find alternative routes.

Exhibit 5-9 Example Prohibited Left Turns by Time of Day

5.1.3.6 Left-Turn Phase Sequence

Regardless of the type of left-turn phasing that is applied, it may be advantageous under certain conditions to change the sequence in which left-turn phases are served (relative to the through phases). There are three sequence options available: lead-lead, lag-lag, and lead-lag.

The most common left-turn phase sequence is the ***lead-lead sequence***, which starts opposing left-turn phases prior to the through phases (see Exhibit 5-6 for an example). The advantages of this sequence option include the following:

- Users react quickly to the leading green arrow indication.
- It minimizes conflicts between left-turn and through movements on the same approach, especially when the left-turn volume exceeds the available storage bay (or no left-turn lane is provided).
- It gives unused time to the through movements.

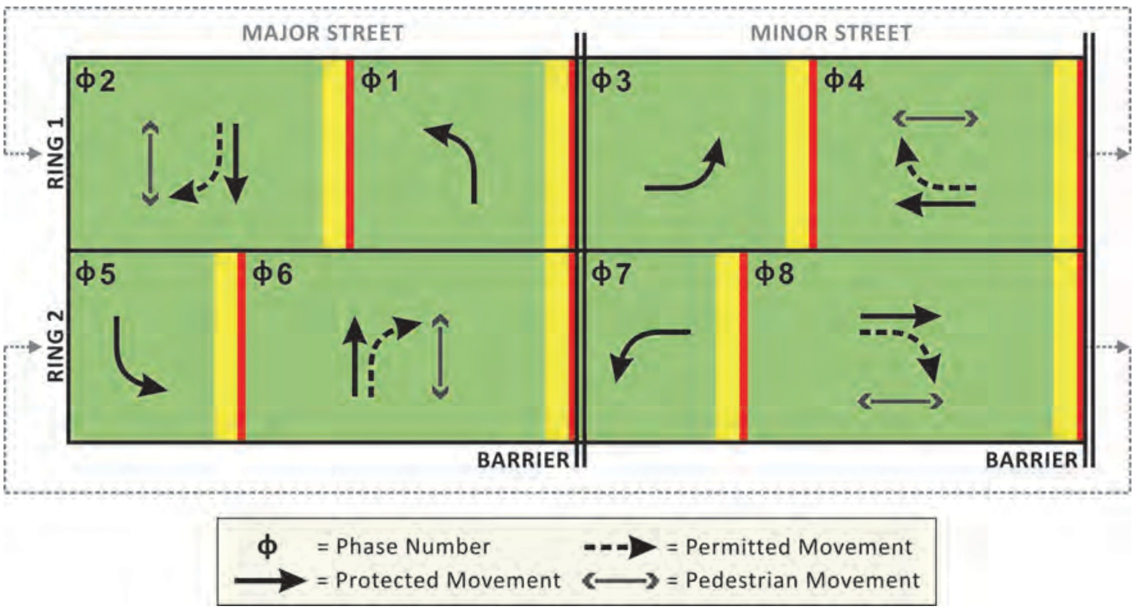
The ***lag-lag sequence***, on the other hand, serves opposing left-turn phases after the through phases. This sequence is most often used in coordinated systems with closely spaced signals (e.g., diamond interchange). Lagging the left-turn phases can have operational benefits when there is an unopposed, protected-permitted left-turn phase (e.g., at a T-intersection or at the intersection of a two-way street and a one-way street). However, the disadvantages of lagging the left-turn phases include the following:

- Users tend not to react as quickly to the lagging green arrow indication.
- If a left-turn lane does not exist (or is relatively short), queued left-turn vehicles may block the inside through lane during the initial through movement phase.
- When lag-lag phasing is used at an intersection with four approaches, where opposing phases are protected-permitted, a yellow trap could exist. (Note that the yellow trap can be alleviated by using a protected-only left-turn phase or a FYA display.)

Opposing left turns can also run in a ***lead-lag sequence***, where they begin and end at different times relative to the through phases (as shown by Phases 1 and 5 in Exhibit 5-10). This sequence is generally used to accommodate through movement progression in a coordinated signal system (discussed in Chapter 7). Like the lag-lag sequence, a yellow trap may exist if protected-permitted phasing is used and should be considered when programming the intersection. Lead-lag phasing can have operational benefits for the following conditions:

- Where there is inadequate space in the intersection to safely accommodate simultaneous service of opposing left-turn movements. (An appropriate controller configuration should be used to ensure that the left-turn phases never time concurrently.)
- Intersections where the leading left-turn movement is not provided with an exclusive lane (or the available left-turn storage bay is relatively small).

Exhibit 5-10 Ring-and-Barrier Diagram Showing Protected Lead-Lag Left-Turn Phasing



5.1.4 Overlaps

An overlap is a separate output that can use special logic to improve operations.

Overlaps provide a way to operate movement(s) with one or more phases. While overlaps are often incorrectly associated with hard-wiring multiple movements to the same phase in the signal cabinet (e.g., right-turn arrow wired to compatible left-turn phase), true overlaps require their own load switch and accompanying signal timing. Overlaps can vary widely in their capabilities and how they time specific intervals. When using overlaps, bench testing is always the best approach to understanding how the overlap actually times its output.

Overlaps are controlled through parent phases (which typically add phases together) and modifier phases (which typically exclude operation during modifier phases) and allow non-conflicting movements to receive a green according to the overlap rules (as demonstrated by the example in Exhibit 5-11). While the right-turn overlap movement in this example could be wired directly to the left-turn parent phase (and, therefore, incorrectly be called an overlap because it achieves the same result), eliminating the use of an overlap load switch decreases flexibility in the signal timing.

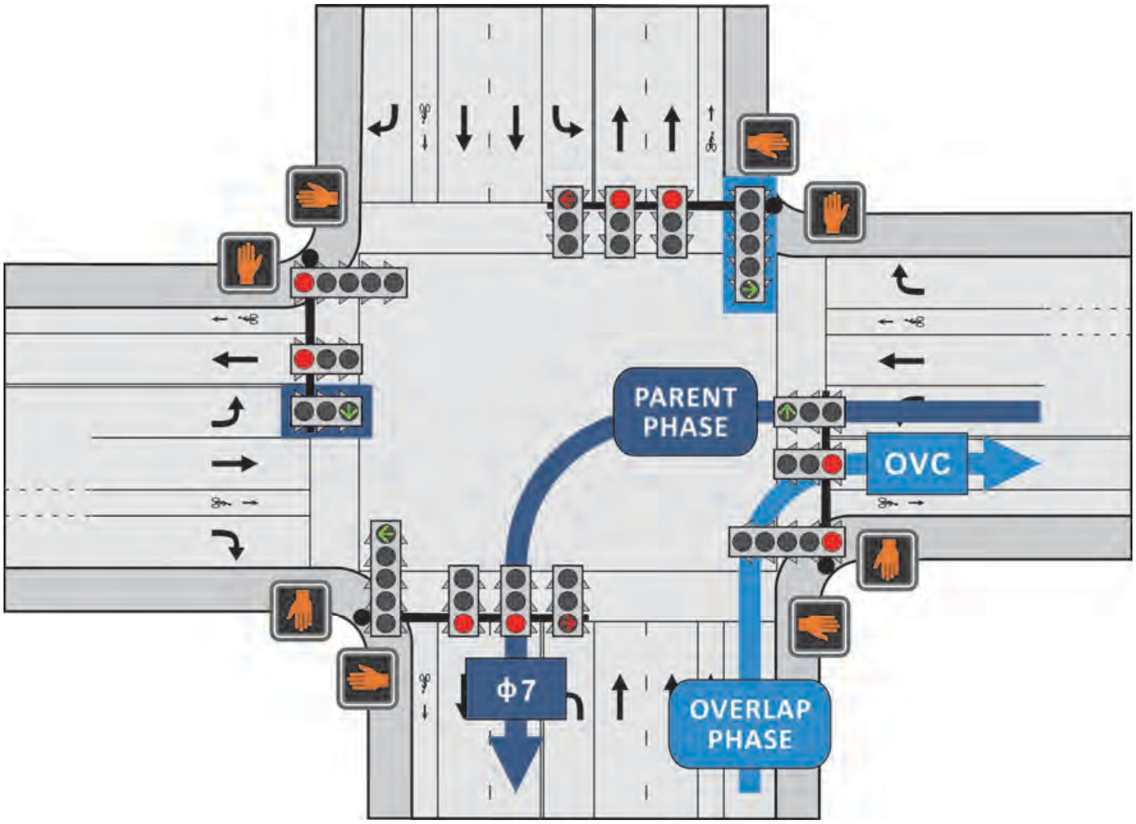


Exhibit 5-11 Example Overlap and Parent Phase

Overlaps can be used to combine phases for any non-conflicting movements, but they are most often used for right-turn movements where exclusive right-turn lanes exist. For right-turn overlaps, the parent phase is typically the compatible left-turn phase. However, it is also possible to operate the right-turn overlaps (subject to the pedestrian conflicts discussed below) with both the adjacent through phase and the compatible left-turn phase. Exhibit 5-12 illustrates a common phase lettering scheme for right-turn overlaps. (Note that overlaps are also sometimes numbered.)

Some traffic signal controllers have a feature that allows a right-turn overlap to be omitted when the conflicting pedestrian phase is active (through the use of a modifier pedestrian phase). This feature is needed if a pedestrian phase is associated with the through vehicular movement. If the right-turn arrow is allowed to turn green at the same time as the through display (and a pedestrian call has been placed), right-turning vehicles will be in conflict with pedestrians receiving a walk indication. An overlap modifier feature will allow the right-turn overlap to have both the compatible left-turn and adjacent through movements as parent phases.

Assigning the conflicting pedestrian phase as the modifier phase (also called Pedestrian Protect, Not Ped Overlap, Conflicting Ped, Negative Pedestrian Phase, or Omit on Green) excludes the right-turn overlap only when there is a pedestrian call on the adjacent through movement. Without the pedestrian modifier, the right-turn overlap must run as a permitted movement with the adjacent through movement to avoid pedestrian conflicts (which is less efficient). Exhibit 5-13 provides a summary of the typical right-turn overlap settings at a standard eight-phase intersection, assuming overlaps are provided for all four right-turn movements. The parent phase designations assume the phase numbering scheme illustrated in Exhibit 5-12.

Exhibit 5-12 Typical Right-Turn Overlap Phase Lettering

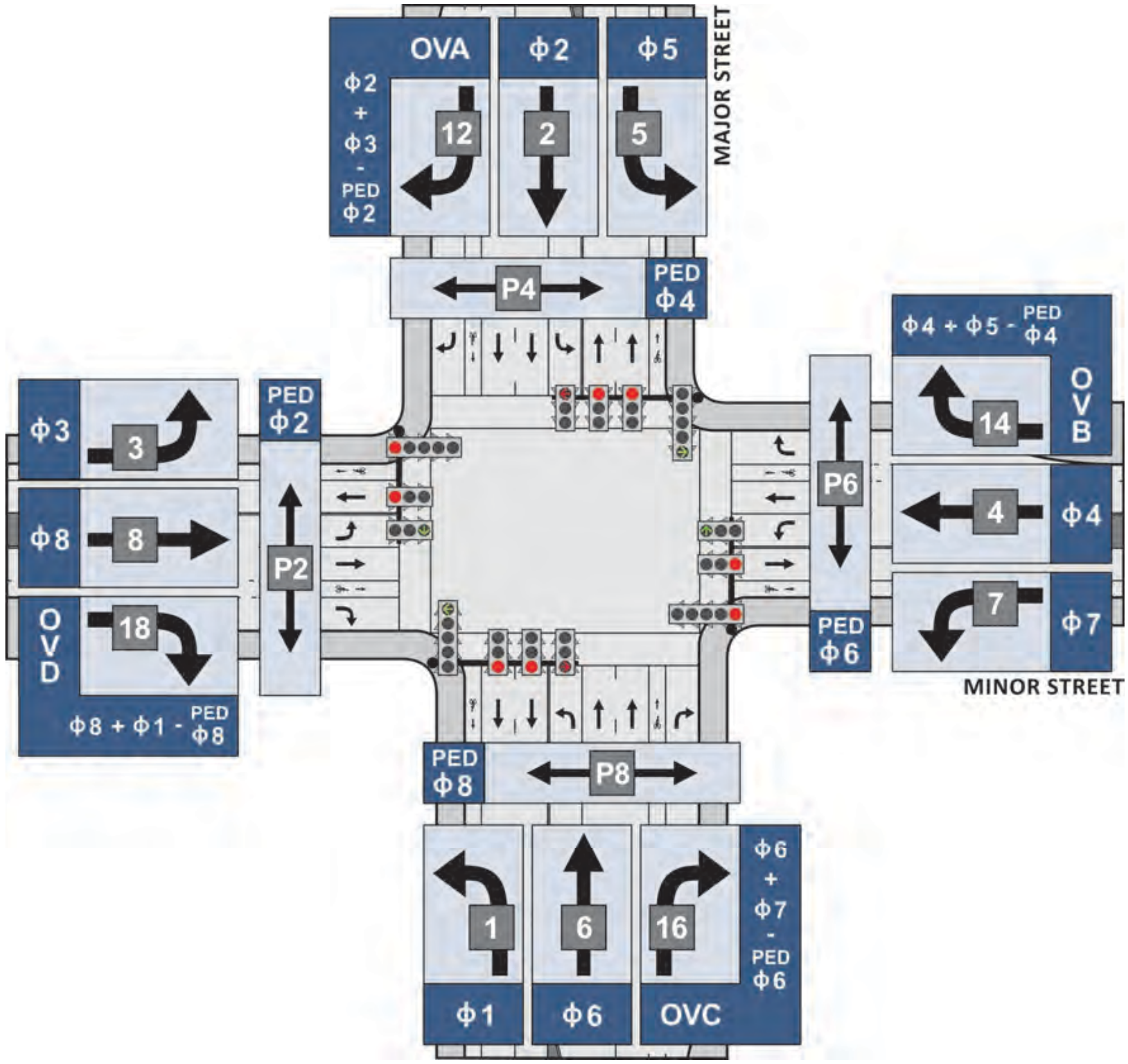


Exhibit 5-13 Typical Right-Turn Overlap Settings

Movement Number	Overlap Letter[1]	Parent Phase	Pedestrian Modifier Phase for Right-Turn Overlap Omit (If Available)
12	A	2* & 3	2P
14	B	4* & 5	4P
16	C	6* & 7	6P
18	D	8* & 1	8P

[1] Agencies may have different overlap assignments based on their preference.
* These phases should not be included as parent phases if a controller feature to omit right-turn overlap with active conflicting pedestrian phases is not available.

Exhibit 5-14 is a ring-and-barrier diagram that shows right-turn overlaps on the major street (during Phases 2 + 3 and Phases 6 + 7). In this example, the right-turn overlaps use a pedestrian modifier function so that the right-turn movements do not conflict with pedestrians utilizing Pedestrian Phases 2 and 6. Assuming that the major street right-turn movements have protected-permitted displays, Overlaps A and C (highlighted in blue) can operate in two different ways during Phases 2 and 6. When a pedestrian phase is called, the adjacent right-turn movement will be permitted (i.e., green ball), and, when there are no pedestrians, the right-turn movement will be protected (i.e., green arrow). Note that if a protected-only display is used with a pedestrian modifier function, the right-turn vehicular movement will be omitted when the conflicting pedestrian phase is called, and a right-turn red arrow will be displayed.

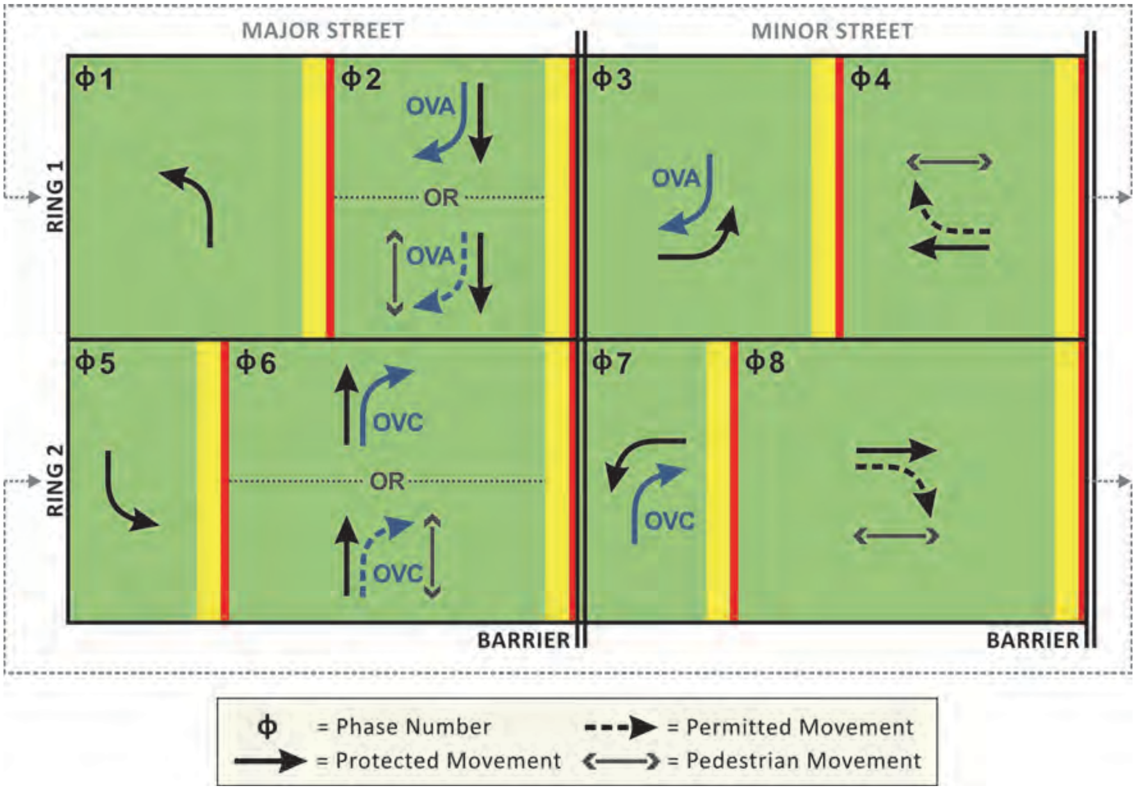
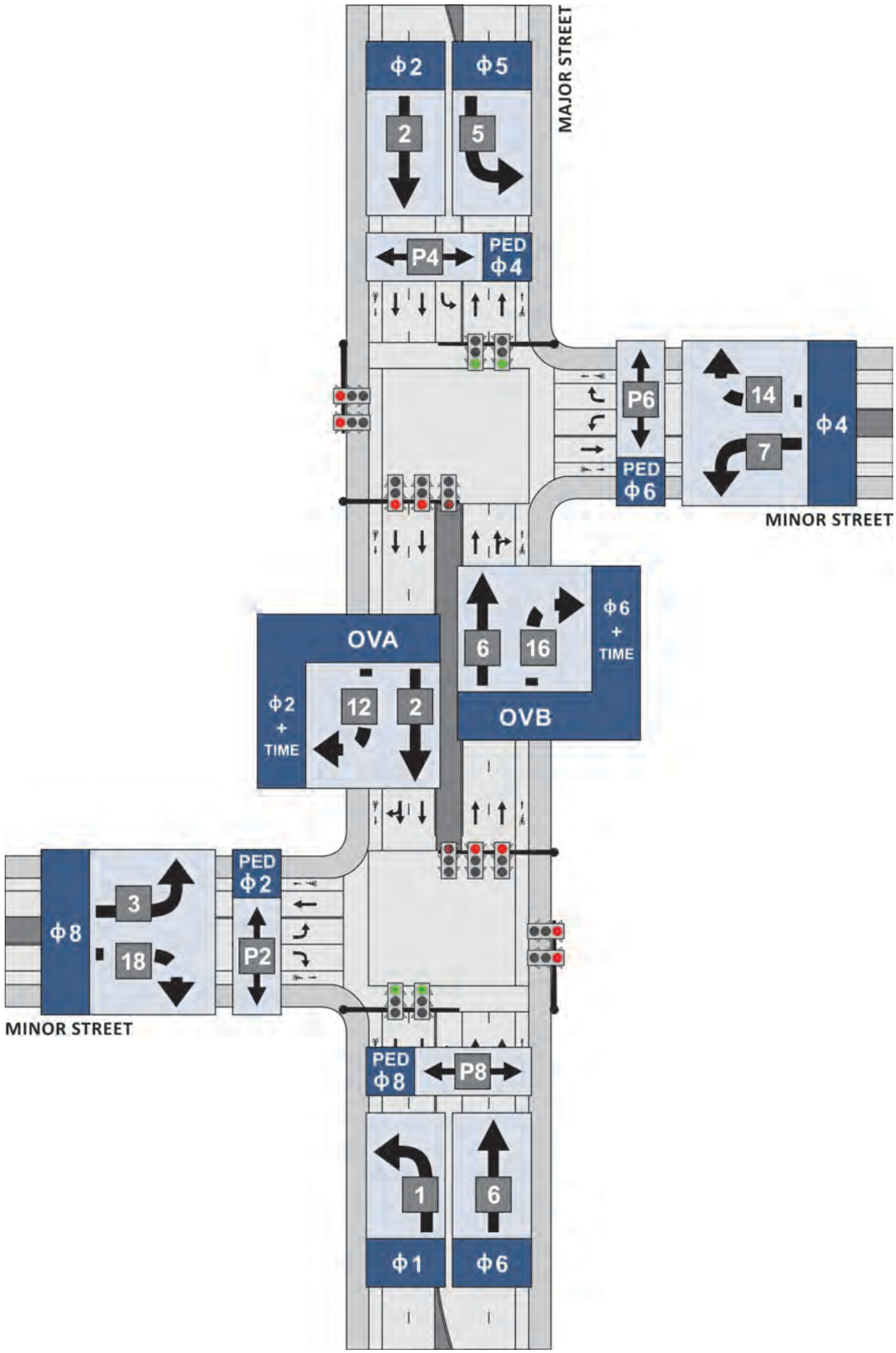


Exhibit 5-14 Ring-and-Barrier Diagram Showing Overlaps (with a Pedestrian Modifier Function)

A trailing overlap is an overlap application that is commonly used at closely spaced intersections (e.g., two closely spaced T-intersections). A trailing overlap continues timing after the parent phase ends. Exhibit 5-15 shows an example intersection where trailing overlaps may be applied. The movements approaching the two-intersection system are assigned phases in a typical fashion (similar to the example in Exhibit 5-1). The movements between the intersections are what differentiates this type of operation. In this case, the southbound movement is assigned to Overlap A (which has Phase 2 as its parent phase), and the northbound movement is assigned to Overlap B (which has Phase 6 as its parent phase). Trailing overlaps are different from standard overlaps because they time with the parent phase ***plus*** a specified amount of time after the parent phase ends.

Exhibit 5-16 shows an example ring-and-barrier diagram associated with the intersection depicted in Exhibit 5-15. Note that the ring-and-barrier diagram illustrates movements for both intersections, using a dashed line in each ring to separate the movements at the northern intersection from those at the southern intersection. In this example, the trailing overlaps (Overlaps A and B) allow for better progression of vehicles between the intersections, reducing the likelihood that a vehicle will get “stuck” between the intersections. The trailing overlaps essentially allow for a set period of time when the space between the intersections is cleared.

Exhibit 5-15 Example of Trailing Overlap Phase Lettering



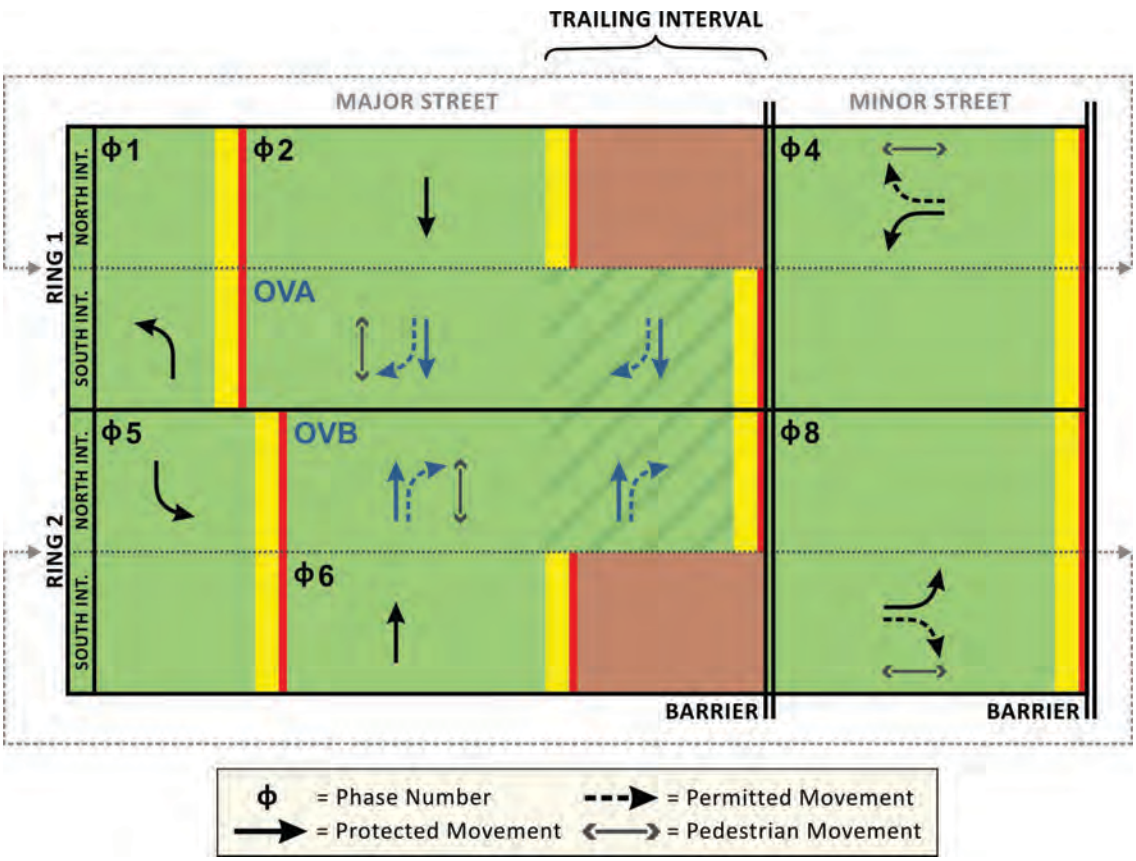


Exhibit 5-16 Ring-and-Barrier Diagram Showing Trailing Overlaps

5.1.5 Detector Assignments

Detector assignments define how the controller will react when detector inputs are received from the field. Basic functions of a detector include calling and/or extending a phase (discussed in detail in Chapter 6). At a fully-actuated intersection, at least one detector is needed to call and extend each phase. The “call” function becomes active only when the phase is not in its green interval, while the “extend” function extends the phase through the use of a timer only during the green interval.

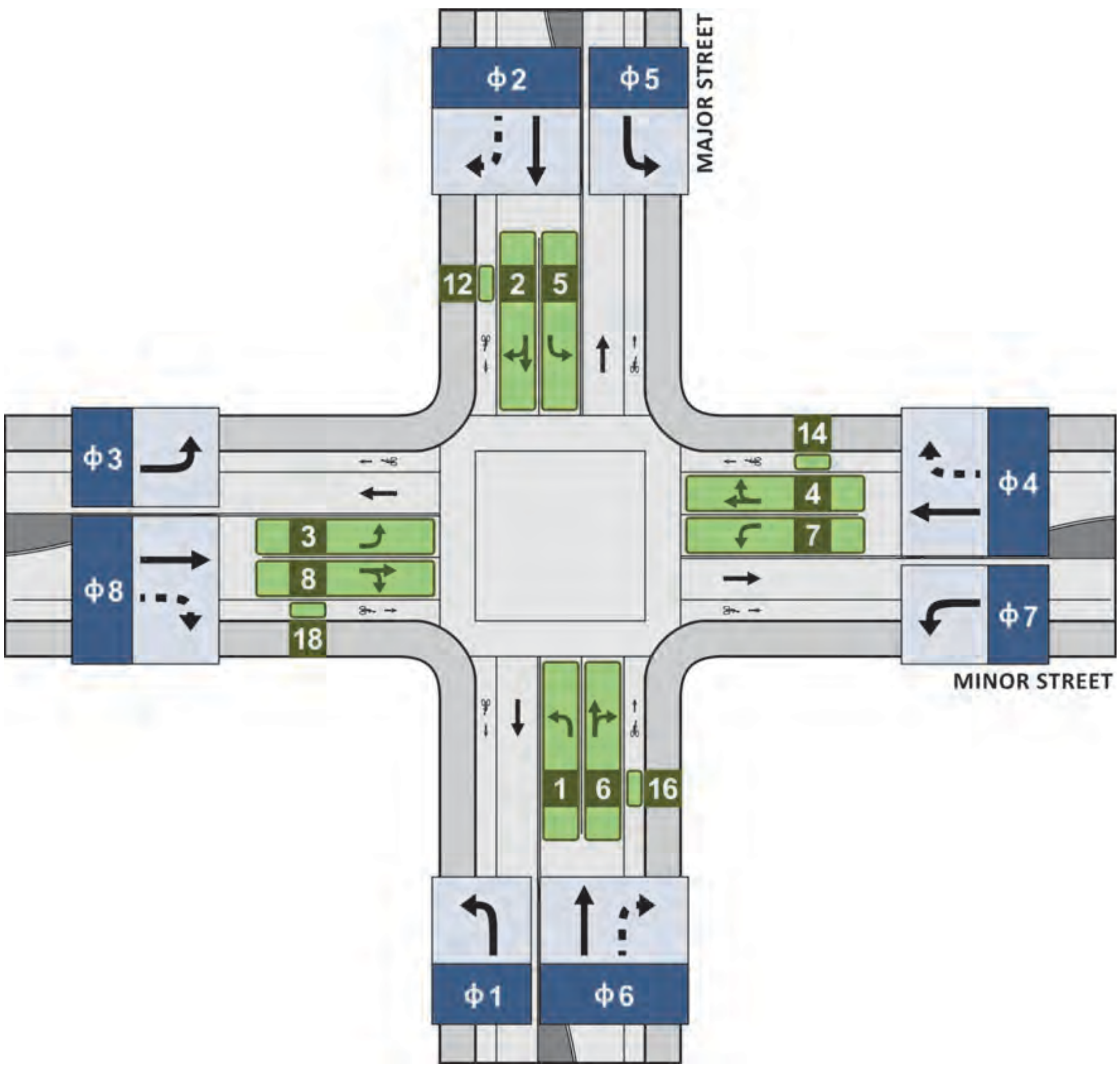
Just as phases and movements are assigned numbers, detectors are also assigned numbers. It is essential to know which number corresponds to each detector so that the detectors can be assigned to the appropriate phase (or phases) in the controller. Exhibit 5-17 summarizes a simple detector assignment at an eight-phase intersection with fully-actuated operations and one detector per lane (as shown in Exhibit 5-18). (Note that most agencies develop a standard that accounts for multiple detectors per phase.) Each detection zone (even those associated with the same phase) will have unique numbers for identification in the signal cabinet. Regardless of the detection numbering scheme that is applied, an agency should use a consistent approach for assigning detector numbers in order to facilitate maintenance and minimize errors.

Exhibit 5-17 Basic Detector Assignment

Phase	Detector Number[1]	Function
Vehicle Phase 1	1	Call & Extend
Vehicle Phase 2	2	Call & Extend
Vehicle Phase 2	12	Call
Vehicle Phase 3	3	Call & Extend
Vehicle Phase 4	4	Call & Extend
Vehicle Phase 4	14	Call
Vehicle Phase 5	5	Call & Extend
Vehicle Phase 6	6	Call & Extend
Vehicle Phase 6	16	Call
Vehicle Phase 7	7	Call & Extend
Vehicle Phase 8	8	Call & Extend
Vehicle Phase 8	18	Call

[1] Detector number will vary depending on field wiring.

Exhibit 5-18 Basic Vehicle Detector Numbering



Most traffic signal controllers do not require additional settings for pedestrian detector assignment. Pedestrian detectors are typically wired such that they are automatically associated with pedestrian phases through a standard input/output configuration (specific to the type of cabinet being used). However, if there is an exclusive pedestrian phase, the standard settings will need to be adjusted for the appropriate phase, consistent with the capabilities of the controller. Changes in phase sequence that differ from the standard phasing scheme of Phases 2, 4, 6, and 8 being the parallel through movements to companion pedestrian phases may also require adjustment of the pedestrian detector inputs.

5.1.6 Load Switch Assignments

In addition to assigning detectors to the phases at an intersection, a practitioner must also associate the displays with each phase. The association between phases/overlaps and vehicular/pedestrian movements is defined by the programmed signal phasing in the controller firmware, the load switch assignment in the traffic signal controller cabinet, and the wiring between each load switch and the signal displays. Most modern traffic signal controller cabinets have default load switch assignments, supported by typical controller firmware, for standard eight-phase intersections. Therefore, additional settings or changes are generally not needed for standard applications. Intersections with eight vehicular phases, four pedestrian phases, and four overlap movements might have the default load switch assignment shown in Exhibit 5-19. Other default settings would be determined by the controller firmware version.

Load switches each have three outputs that are generally associated with red, yellow, and green indications. Pedestrian displays use the same type of load switch as vehicle displays, but they only use the "green" output for walk and the "red" output for don't walk. Flashing don't walk is the "red" output flashed on and off every half second. For indications with more than three outputs, such as an FYA display, unused load switch outputs can be reassigned. For an FYA display, the flashing yellow indication is typically run off an unused load switch output (e.g., pedestrian phase "yellow") or an unused load switch (e.g., an entirely unused overlap or phase). The use of the pedestrian "yellow" output is the most efficient use of load switches, but it is not possible with some controllers, and may not be compatible with some signal monitors. The practitioner will have to verify compatibility before reassigning load switches.

Signal monitors are responsible for monitoring conflicting phases and conflicting indications in a single display. (More information is available in Chapter 4.)

Exhibit 5-19 Typical Load Switch Assignment

Phase / Overlap	Load Switch Number
Vehicle Phase 1	1
Vehicle Phase 2	2
Vehicle Phase 3	3
Vehicle Phase 4	4
Vehicle Phase 5	5
Vehicle Phase 6	6
Vehicle Phase 7	7
Vehicle Phase 8	8
Overlap A	9
Overlap B	10
Overlap C	11
Overlap D	12
Pedestrian Phase 2 + FYA1*	13
Pedestrian Phase 4 + FYA3*	14
Pedestrian Phase 6 + FYA5*	15
Pedestrian Phase 8 + FYA7*	16

* FYA displays often use the pedestrian phase yellow load switch.

5.2 CRITICAL MOVEMENT ANALYSIS

Although more complex methodologies exist, critical movement analysis may be the most appropriate for certain situations.

An important principle behind effective signal timing plans is the relationship between signal timing, phasing, and the capacity of an intersection. A variety of analysis procedures, ranging from simple to complex, can be used to evaluate signalized intersection performance. Critical movement analysis is a simplified technique that has broad applications for estimating phasing needs and signal timing parameters. This method allows a practitioner to identify the critical phase pairs at an intersection, calculate the critical volume, and approximate the required cycle length.

The method is generally simple enough to be conducted by hand (making it convenient for use in the field), although some of the more complicated refinements are aided considerably by the use of a simple spreadsheet. In many cases (such as signals at new streets), critical movement analysis is the appropriate level of analysis given the lack of accuracy in future volume forecasts. It can also provide a first-cut reasonableness check of software results, and, therefore, is an important procedure to understand as a signal timing practitioner.

The following sections demonstrate the critical movement analysis procedure using two simple examples. While both examples assume the same lane configuration and vehicular volumes at the intersection, they assume different phase sequences. The first example demonstrates the critical movement analysis procedure using a typical eight-phase intersection with protected left turns on all approaches (Example #1 shown in Exhibit 5-20), while the second example uses an intersection with split phasing on the minor street (Example #2 shown in Exhibit 5-21).

Exhibit 5-20 Example #1: Critical Movement Analysis Phases and Volumes

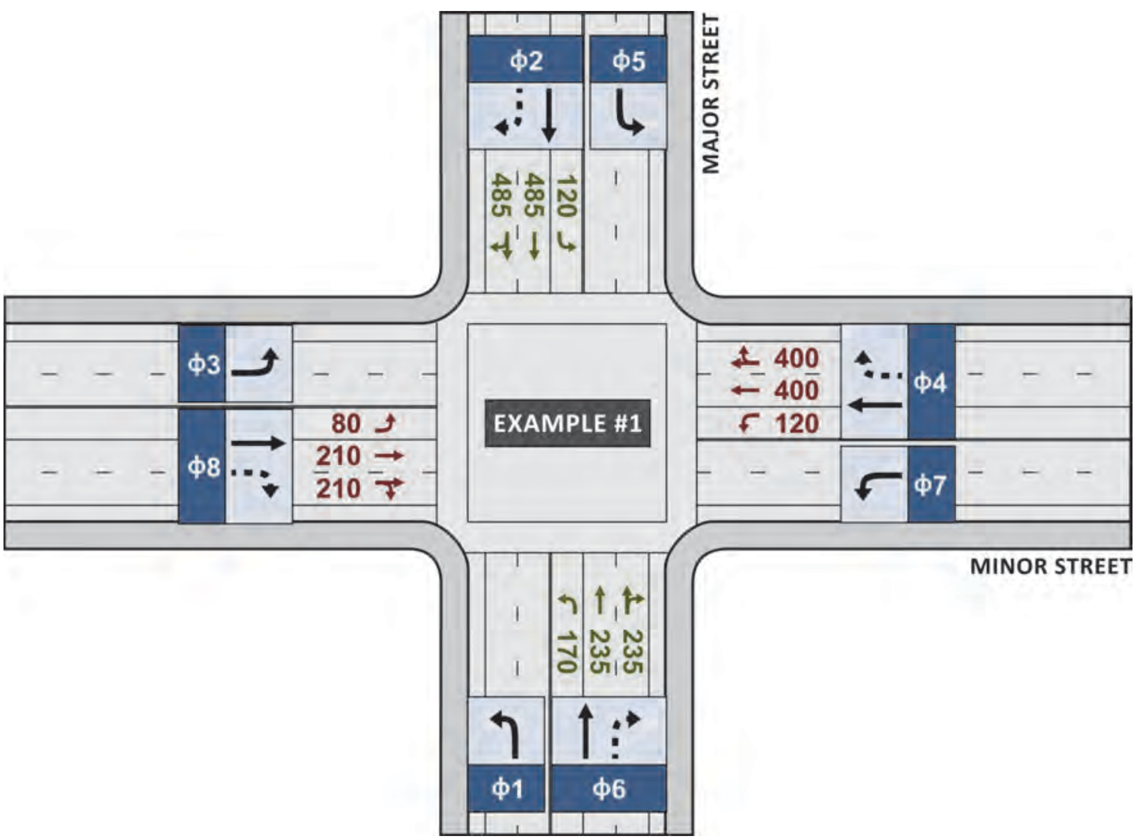


Exhibit 5-21 Example #2: Critical Movement Analysis Phases and Volumes

5.2.1 Step 1: Record Demand Volumes

The first step in the critical movement analysis procedure is determining peak period volumes in each lane at the intersection (as shown in Exhibit 5-20 and Exhibit 5-21). It should be noted that the results of an analysis are highly dependent on the accuracy and also the variability in traffic volumes. Care should be taken in interpreting the results, and the interpretation should be based on an understanding of the fundamental quality of the data. There are several adjustments that should be made to the volumes depending on site-specific conditions:

- ***Peak-Hour Adjustment.*** If hourly traffic volumes are being used, they are usually adjusted to reflect the peak 15-minute period.
- ***Heavy-Vehicle Adjustment.*** The practitioner should always evaluate the proportion of heavy vehicles and make adjustments to account for larger vehicles if appropriate. Modest heavy-vehicle volumes are assumed in this example, so no adjustments have been made. If significant numbers have been counted, trucks and buses can be considered as two vehicles each in the analysis.
- ***Lane Imbalance Adjustment.*** Volumes should be adjusted to account for site-specific conditions if significant lane imbalances exist due to downstream destinations. In this example, volumes are evenly distributed among the lanes.

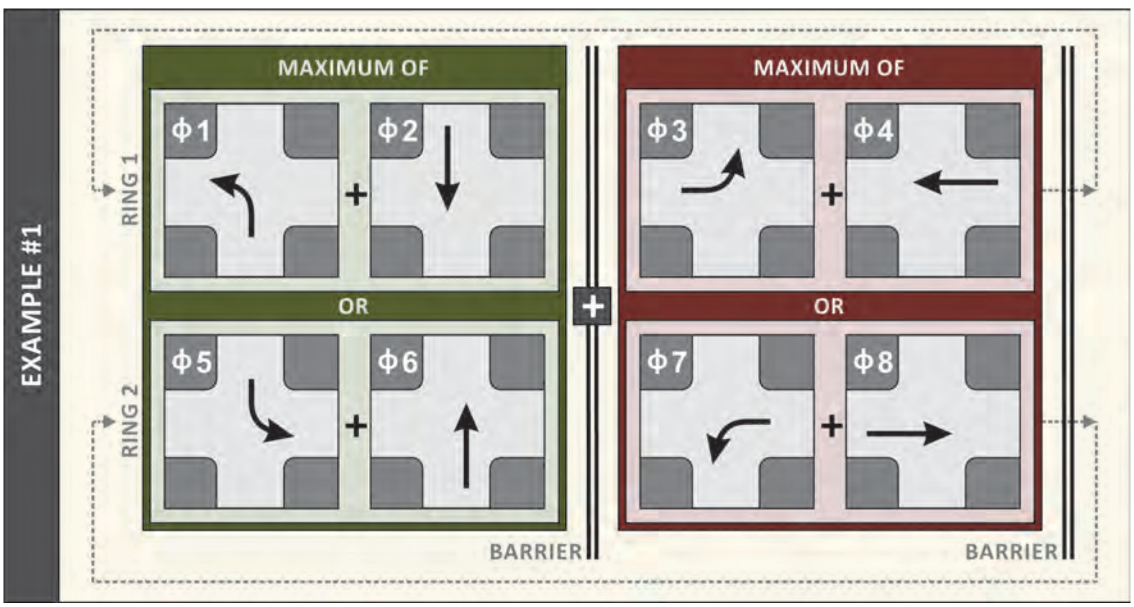
5.2.2 Step 2: Determine Critical Phase Pairs

The cycle length needed to accommodate all of the vehicles at an intersection can be estimated by identifying the movements that require the most time (the critical movements). In the second step of the critical movement analysis procedure, the vehicular volumes associated with conflicting phases (generally a left-turn phase and opposing through movement phase) are used to determine the critical phase pairs.

5.2.2.1 Example #1: Protected Left-Turn Phasing

The volume sums for all sets of conflicting phases are compared between the barriers (in Example #1, Phases 1 + 2, Phases 3 + 4, Phases 5 + 6, and Phases 7 + 8), as depicted in the ring-and-barrier diagram in Exhibit 5-22. This comparison allows a practitioner to identify the critical movements (and associated critical volumes) for the major street and minor street. In Example #1, the amount of time given to the major street will be dictated by whichever conflicting set of phases (Phases 1 + 2 or Phases 5 + 6) has the most volume. Similarly, the amount of time given to the minor street will be dictated by whichever conflicting set of phases (Phases 3 + 4 or Phases 7 + 8) has the most volume. This will ensure that the cycle length is based on the movements with the most volume (that thus require the most time). Exhibit 5-23 shows the volumes from Exhibit 5-20 assigned to each phase in the ring-and-barrier diagram.

Exhibit 5-22 Example #1: Conflicting Phases



Note that all of the through movements are associated with two lanes. The highest volume in any lane associated with the same phase should be used to identify the critical movements. Typically, through lanes will experience fairly uniform volumes, but there are exceptions to this, such as at locations near freeway interchanges. As an example, note that there are two through lanes assigned to Phase 2 in Exhibit 5-20. If the figure showed that there were 385 through vehicles in the inside lane and 585 through vehicles in the outside lane, the Phase 2 volume used for the conflicting phase volume calculation would be 585 vehicles.

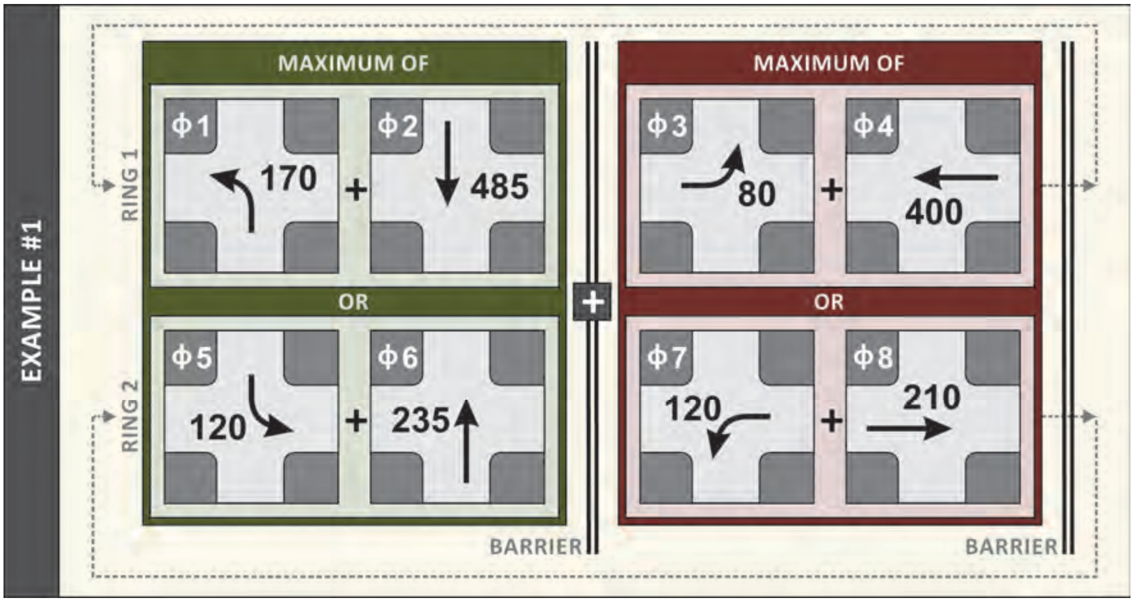


Exhibit 5-23 Example #1: Volumes by Phase

In Example #1, Phases 1 and 2 have more conflicting volume than Phases 5 and 6 (as depicted in Exhibit 5-24). In other words, the volume proceeding through the intersection during Phases 5 and 6 will be able to clear the intersection before the volume during Phases 1 and 2, so Phases 1 and 2 are the major street critical movements. Phases 3 and 4 have more conflicting volume than Phases 7 and 8, so they are the minor street critical movements.

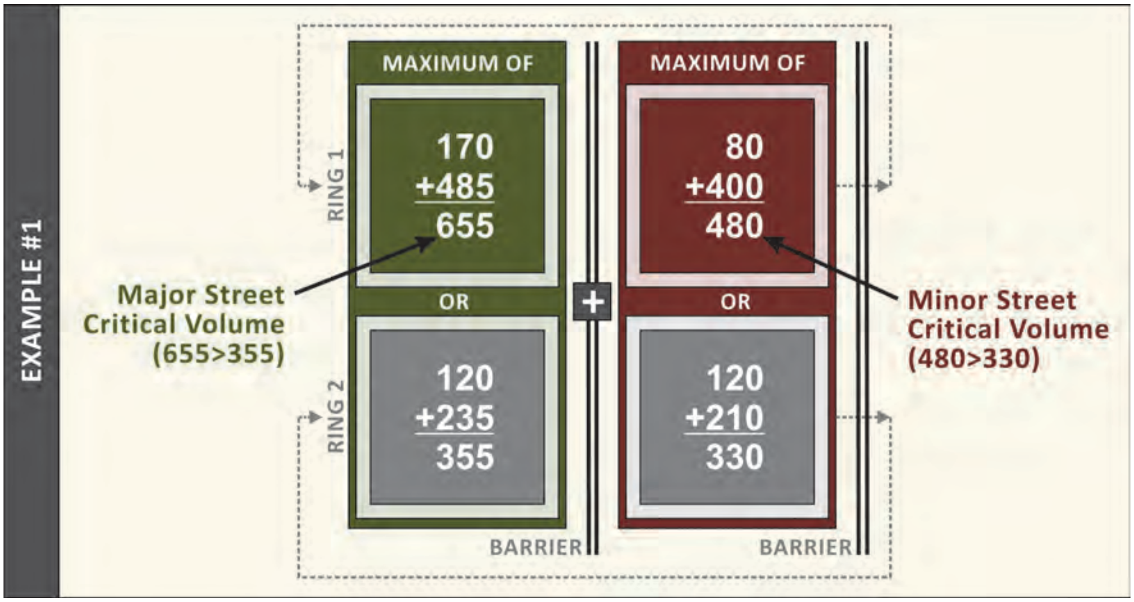
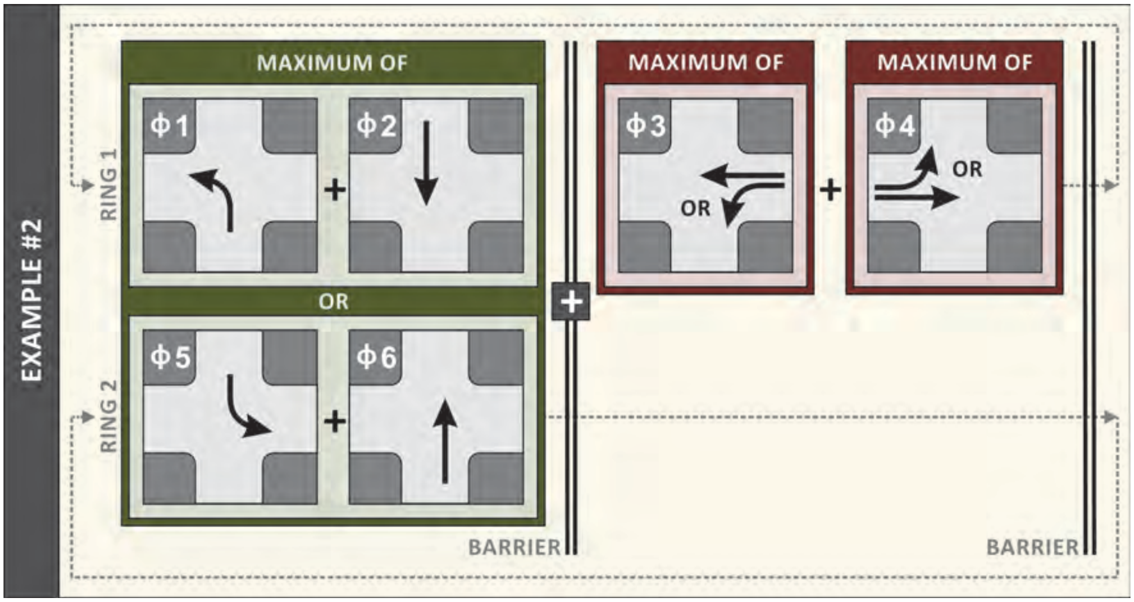


Exhibit 5-24 Example #1: Major and Minor Street Critical Volumes

5.2.2.2 Example #2: Split Phasing on the Minor Street

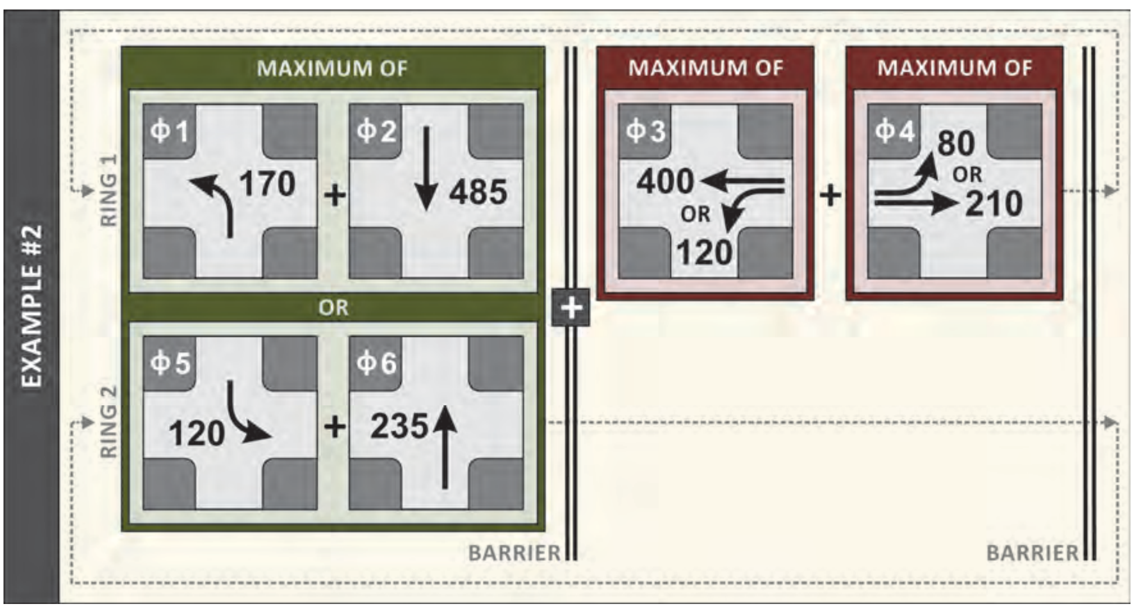
Note that different phase sequences may change which phases are considered in the critical movement analysis. In Example #2, split phasing at the intersection allows all movements from one minor street approach to proceed through the intersection at the same time (Phase 3), followed by all the movements from the other minor street approach (Phase 4). Exhibit 5-25 shows that the critical volume for the minor street will be determined by summing the greatest volume from each minor street approach. As a result of the split phasing, the minor street demand will likely require more time to serve than if the minor street were using permitted or protected left-turn phasing.

Exhibit 5-25 Example #2: Conflicting Phases



In Example #2, the major street critical movements will be identified using the same method explained in Example #1 (whichever conflicting set of phases has the most volume between Phases 1 + 2 or Phases 5 + 6). However, determining the critical movements on the minor street will involve a slightly different computation (as depicted in Exhibit 5-26).

Exhibit 5-26 Example #2: Volumes by Phase



The minor street is using only two phases (Phases 3 and 4) because all of the vehicles from one approach will proceed through the intersection before all of the vehicles from the other approach will proceed through the intersection. The two phases are in the same ring because they cannot run at the same time. (There are other ways to accomplish this, but this is the most common.) The amount of time required for the minor street will be dictated by the sum of the highest lane volumes on each minor street approach.

As explained previously, the major street critical volume will be the same in Examples #1 and #2. The minor street critical volume in Example #2, however, is determined by adding the highest lane volume from each minor street approach (as depicted in Exhibit 5-27). The through volumes are higher than the left-turn volumes for both phases (Phases 3 and 4), so the through volumes will determine how much time is required for the minor street. During each minor street phase, left-turning vehicles should be able to clear the intersection before the through vehicles. Phases 3 and 4 run separately under split phasing, so the minor street critical volume is the sum of the highest lane volumes for Phases 3 and 4.

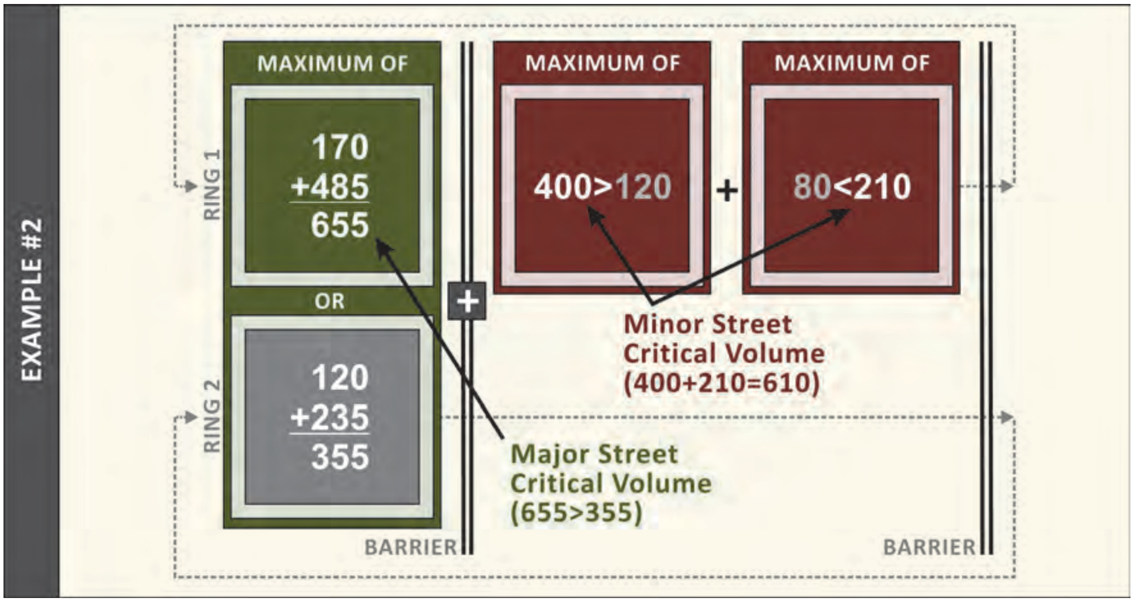


Exhibit 5-27 Example #2: Major and Minor Street Critical Volumes

5.2.3 Step 3: Calculate the Critical Volume

The third step in the critical movement analysis procedure is determining the total demand on the intersection. The critical volume on the major street and the critical volume on the minor street should be summed to determine the critical intersection volume (as shown in Exhibit 5-28 for Example #1 and Exhibit 5-29 for Example #2).

The critical intersection volume is the highest number of vehicles that must be accommodated at the intersection based on the phase sequence and demand on each movement. Note that the critical volume for Example #2 is higher than the critical volume for Example #1, despite the same lane volumes from Exhibit 5-20 and Exhibit 5-21. More time will be required for the minor street under split phasing because vehicles from each minor street approach must proceed through the intersection separately. There is less efficiency with this phase order.

Exhibit 5-28 Example #1: Critical Volume

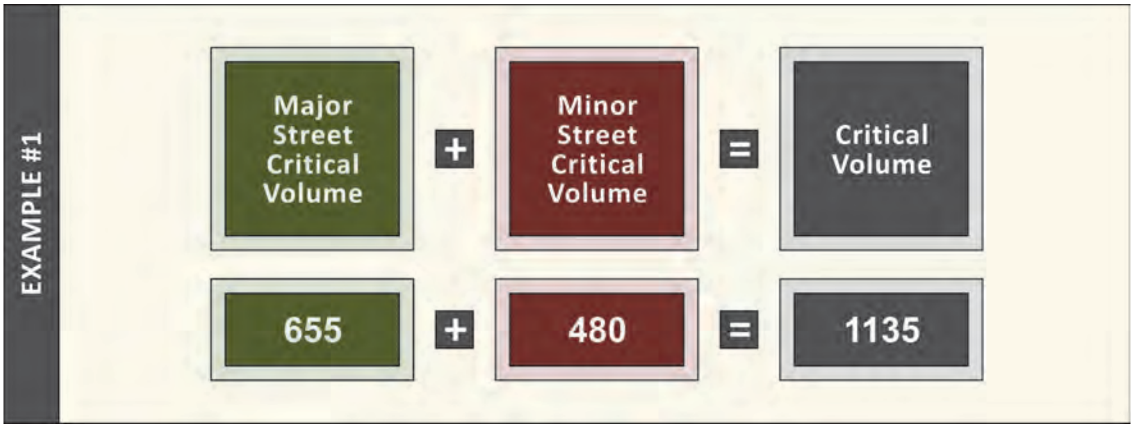
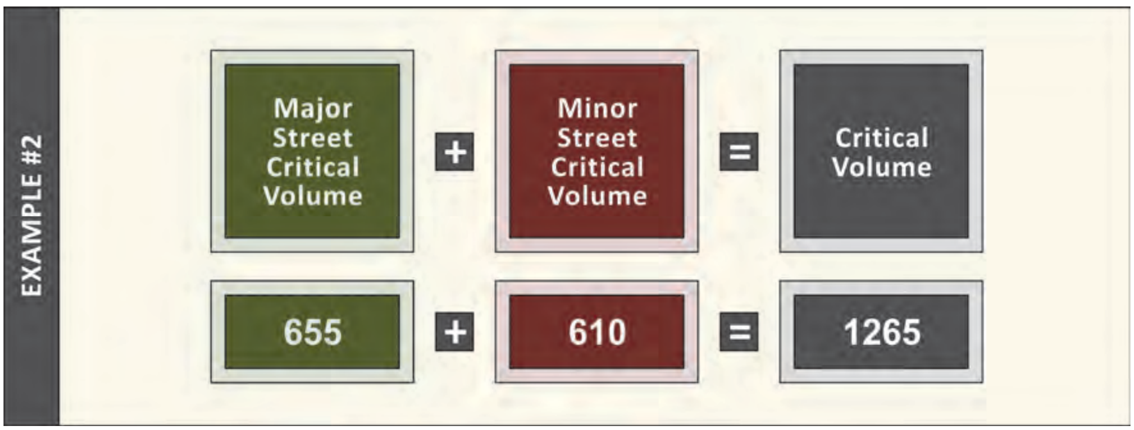


Exhibit 5-29 Example #2: Critical Volume



5.2.4 Step 4: Estimate the Cycle Length

The critical volume at an intersection can be used to estimate the required cycle length. Exhibit 5-30 shows the number of vehicles that can be accomodated per cycle and per hour for an intersection with eight phases (based on cycle lengths between 60 seconds and 120 seconds). Effective green accounts for start-up lost time as well as vehicles moving during a portion of the yellow. Start-up lost time is the additional time consumed by the first few vehicles in a queue above and beyond the saturation headway, because of the need to react to the green indication and accelerate. Effective green is approximately equal to actual green in typical applications (i.e., no excessive start-up lost time or excessively long clearance intervals). The calculations assume an intersection capacity of 1,400 vehicles and 5 seconds of lost time per phase. In Example #1 (with eight phases and a critical volume of 1,135 vehicles), a cycle length of 110 seconds would accommodate the intersection demand based on a relatively conservative set of assumptions. (These conservative assumptions will generally provide an upper bound on cycle length.)

The signal timing practitioner should use judgment when selecting signal timing values based on simple demand analysis, as the computed values may not be consistent with operational objectives. Intersection demand is just one factor that should affect the cycle length that is chosen. For example, the average segment length, street classification, left-turn phasing, pedestrian phasing, and objectives determined for the intersection can also influence the optimum cycle length. If block lengths are long, a longer cycle length may be appropriate to favor major street traffic. If blocks are short or pedestrian volumes high, shorter cycle lengths may be more appropriate.

Exhibit 5-30 Estimated Cycle Lengths Based on Critical Volume (Eight-Phase Intersection)

Cycle Length (Seconds)	Number of Cycles Per Hour	Lost Time Per Cycle (Seconds)[1]	Effective Green Time Per Cycle (Seconds)	Number of Vehicles Per Cycle[2]	Maximum Number of Vehicles Per Hour[2]
60	60	20	40	16	933
70	51	20	50	19	1000
80	45	20	60	23	1050
90	40	20	70	27	1089
100	36	20	80	31	1120
110	33	20	90	35	1145
120	30	20	100	39	1167

[1] This lost time assumes that the intersection is operating with eight phases (four in each ring) with 5 seconds of lost time per phase. The lost time will be less at an intersection with fewer phases.
[2] The number of vehicles that can be accommodated under the various cycle lengths was calculated assuming a headway of 2.5 seconds per vehicle, which is generally conservative for urban/suburban environments.

It should be noted that the longer the green time per cycle, often the larger the vehicle headways (i.e., lower vehicle density) later in the phase. This can result in less efficient use of green time at the intersection, which is akin to lost time when aggregated. Longer cycle lengths are an effective way to move large numbers of vehicles at an intersection if the desired headways are maintained during the entire green time. Queue storage, upstream and downstream bottlenecks, and density of vehicles all play into effective movement of vehicles through a signalized intersection, and the needs of all users should be considered when selecting a cycle length for implementation.

5.3 ROLE OF SOFTWARE IN SIGNAL TIMING

Computer-based tools are available to calculate and evaluate signal timing. It is important to recognize their capabilities and limitations, and it is recommended that practitioners develop a thorough understanding of the selected computer programs, their uses, and how they relate to field conditions. Using the optimize feature of a software product that does not use an agency's operational objectives will not produce the desired outcome. In other words, a sophisticated tool in the hands of an inexperienced analyst may not produce a satisfactory result.

Operating environment and timing objectives should be considered before a software tool is selected.

5.3.1 Types of Software Models

In general, there are two types of models that can be developed using software tools: (1) deterministic or equation-based models and (2) microscopic simulation models. These two types of models assess different levels of detail related to traffic operations, and should be applied to signal timing projects based on the traffic conditions and complexity of the network. It should be noted that all models are only as good as the inputs, as well as their imbedded parameters. Saturation flow rates, regardless of whether they are fixed or adjustable, are generally assumed to be constant, which is not necessarily the case as cycle length increases or access disrupts flow.

Deterministic or equation-based models take multiple intersections into account, unlike isolated intersection methods.

5.3.1.1 Deterministic or Equation-Based Models

For many practitioners, deterministic or equation-based models are the models of choice when developing signal timing plans, particularly for coordinated systems. Deterministic or equation-based models are different than individual intersection methods (such as critical movement analysis) because they account for the arrival and departure of vehicles from one intersection to the next, otherwise known as the system effect. Traffic progression is treated explicitly through the use of time-space diagrams (explained in Chapter 7) or platoon progression techniques. Some models are also capable of deterministically estimating the effect of detection parameters for both vehicles and pedestrians. They typically use a combination of scaling factors for vehicle demand and the presence or absence of vehicle and/or pedestrian calls.

A key feature of these signal timing models is the explicit effort to “optimize” signal timing to achieve a particular performance measure (such as delay), which may or may not tie into the desired outcome. To accomplish its optimization, each model uses some type of algorithm to test a variety of combinations of cycle length, splits, and offsets to achieve a calculated value of one or more performance indices, and then attempts to find an optimal value for those performance indices. Most deterministic or equation-based models use elements of HCM procedures to estimate certain parameters, such as saturation flow rate and delay. Saturation flow rate is the hourly rate at which vehicles can traverse an intersection approach under prevailing conditions, assuming a constant green indication and no lost time.

It is important to understand the appropriateness of the model as it relates to the operational objectives that the practitioner wants to achieve. If the operational objective is smooth arterial flow with minimal stops, then the output from a delay minimization software tool may need to be manually adjusted to obtain values that are appropriate for the operational objective. For example, increasing the cycle length slightly may not correspond to the minimum possible delay (as determined by the software tool), but may significantly reduce the number of stops.

These deterministic or equation-based models are frequently sufficient for most signal timing applications. However, they can lose validity in cases where demand exceeds capacity or where the queues from one intersection interact with the operation of an adjacent intersection. In congested conditions, reducing the cycle length may actually increase throughput by reducing queues which block movements at other intersections. Therefore, locations with closely spaced intersections or with intersections exceeding capacity may not be well served by these types of deterministic models. In these cases, it may be necessary to use microscopic simulation models to obtain more realistic assessments of signal timing effects.

5.3.1.2 Microscopic Simulation

In the context of signal timing, microscopic simulation models can be thought of as an advanced evaluation tool. They can estimate car-following behaviors for vehicles and have the ability to model other users at a signalized intersection (including pedestrians, bicycles, and transit). Recent advances in technology also allow direct linkages between simulation models and either actual signal controllers or software emulations of those controllers, known as hardware-in-the-loop (HITL) and software-in-the-loop (SITL), respectively. These in-the-loop simulations allow actual controllers and/or their algorithms to replace the approximation used in simulation models, more accurately reflecting how a controller will operate.

The simulation model is used to generate traffic flows and send vehicle and pedestrian calls to one or more controllers (based on the detection design implemented in the simulation). The controller receives the calls as if it was operating in the field and uses its own internal algorithms to set signal indications based on the calls received and the implemented signal timing. The signal displays are passed back to the simulation model to which traffic responds. This type of analysis is particularly effective for modeling special controller features such as transit signal priority and railroad preemption, because these features are controller firmware (logic in hardware) specific.

Microscopic simulation models are advanced tools capable of approximating more signal timing controls than deterministic or equation-based models.

The *Traffic Analysis Toolbox* (*3*) describes simulation models as particularly useful for the following applications:

- To evaluate signals incorporating actuated-coordinated operations. The simulation program (if it includes realistic controller features and is coded correctly) may provide a more realistic assessment of the effectiveness of the actuated controller within the section being evaluated.
- To confirm the likely presence of queue spillback between intersections.
- To evaluate special features, such as transit priority (although generic transit signal priority models may not reflect the type of transit priority actually in use).
- To evaluate system performance in the presence of saturated conditions.
- To evaluate the effectiveness of manual adjustments to signal timing.
- To evaluate fuel consumption and emissions or system travel time resulting from a given set of signal timing.
- To evaluate various travel modes (i.e., pedestrian and bicycle traffic).
- To demonstrate improvements to public officials, as they paint a clearer picture than numbers alone.

This last application refers to the use of simulation models for their animated graphical outputs, which approximate aerial or 3-D views of the simulated network. These animated outputs permit visual assessment of system performance, a capability that cannot be duplicated by any other form of output, including time-space diagrams. Care must be used when reviewing animation from simulation models, as they should be reflective of an average run (which can be selected only after a review of the results).

When running simulation analyses to compare alternatives, the practitioner should ensure that the resolution of the model is sufficient to model the type of control anticipated. A simulation model, for example, typically requires precision on the order of a tenth of a second to accurately replicate gap detection. Therefore, the use of coarser resolution settings (or a model with only 1-second resolution) to speed up simulation run time may yield incorrect results.

5.3.2 Software Considerations

There are several items that should be considered when choosing which type of software tool to apply or whether to use a software application at all. The detector settings, saturated conditions of the area, user priority assumptions, unique network

A good rule of thumb is to have a sense for the expected answer (based on field knowledge or a quick critical movement analysis) to check if the software results are reasonable.

features, and required field calibration should be taken into account when choosing a software tool.

5.3.2.1 Detector Settings

Accurate detector locations, sizes, and timing parameters—often overlooked in software—are very important for creating accurate simulation results. A practitioner should determine the existing detector locations and settings (or plans for future implementation), as well as understand typical practice (i.e., standard detector layout plans) in order to build simulation models that are reliable. Software algorithms are subject to detector inputs, which can artificially gap out early or extend unrealistically, creating inaccurate simulation results. Care should be taken to match detector locations, sizes, and timings accurately.

5.3.2.2 Saturated Conditions

Saturated conditions require careful distribution of green time that balances queue buildup and meters incoming demand. In addition, queued vehicles should be stored at locations where they will not impede other traffic. Because under-saturated flow tends to be stable (similar from one cycle to the next), traffic signal timing software can often calculate optimal signal timing more easily under those conditions. However, many tools automatically assume under-saturated conditions, which can make evaluating oversaturated conditions difficult or produce erroneous results. As noted in the previous section, ***near, at, or oversaturated conditions may best be modeled by an experienced practitioner using a microscopic simulation model***. Strategies related to oversaturated operations are discussed in Chapter 12.

5.3.2.3 User Priority Assumptions

Many software tools make assumptions about user priorities without input from the practitioner. Review of user priorities should be completed when performing signal timing evaluations in order to assess whether the default software priorities match the practitioner-defined priorities. Adjustments may need to be made to the signal timing “optimized” by the software.

5.3.2.4 Unique Network Features

Some unique network features, such as transit priority or a railroad, may warrant the use of a specific software tool. Simple analysis techniques, such as critical movement analysis, will not be able to evaluate the effect of those types of features on vehicle progression or intersection capacity.

5.3.2.5 Field Observation and Calibration

Field observations should be compared with traffic operation results for each time period in the model to validate that traffic volumes are correct. Then, if necessary, the base networks for each of the time periods should be calibrated using travel time, delay, and queue data collected from the field. Parameters within the models that can be adjusted to calibrate the existing base networks with actual field conditions include saturation flow rates, right-turning-vehicles-on-red, and lane utilization. A review of the calibrated model should be performed prior to moving forward with the timing plan development. It should be noted, however, that field observation is only a snapshot in time. Modern controllers now have significant capabilities for assessing performance and these capabilities should not be overlooked in modeling operations.

5.3.3 Software Inputs and Outputs

Signal timing plan development typically concludes with the use of a software tool that helps determine appropriate cycle lengths, offsets, and splits for the intersections. In order to use a software tool to assist with the selection of such values, the practitioner first needs to determine some of the basic signal timing parameters. Exhibit 5-31 provides a summary of the sections where typical software inputs and outputs are described in detail throughout this chapter and Chapters 6 and 7.

Using the guidance from Chapter 6, a practitioner should be able to define most of the required software inputs. Chapter 7 provides additional information about the typical software outputs for a coordinated system (cycle lengths, offsets, and splits) and some information about the relationship between basic signal timing parameters and coordination. However, the practitioner should refer to software manuals for specific information about software operations.

Exhibit 5-31
References for Typical Software Inputs and Outputs

	Signal Timing Parameter	Chapter 5 Reference	Chapter 6 Reference	Chapter 7 Reference
Inputs	Phase Sequence	*5.1.1 Movement and Phase Numbering* *5.1.2 Ring-and-Barrier Concept* *5.1.3 Left-Turn Phasing* *5.1.4 Overlaps*		*7.2.6 Bandwidth* *7.6.1 Phase Sequence*
Inputs	Minimum Green		*6.1.3 Minimum Green*	
Inputs	Maximum Green		*6.1.4 Maximum Green*	*7.4.3 Splits Guidance* *7.4.4 Force-Offs Guidance*
Inputs	Yellow Change		*6.1.1 Yellow Change*	
Inputs	Red Clearance		*6.1.2 Red Clearance*	
Inputs	Leading/Lagging Left Turns	*5.1.3 Left-Turn Phasing*		*7.2.6 Bandwidth* *7.6.1 Phase Sequence*
Inputs	Passage Time		*6.1.5 Passage Time (Unit Extension or Gap Time)*	
Inputs	Minimum Gap		*6.1.5 Passage Time (Unit Extension or Gap Time)*	
Inputs	Time Before Reduction		*6.1.5 Passage Time (Unit Extension or Gap Time)*	
Inputs	Time to Reduce		*6.1.5 Passage Time (Unit Extension or Gap Time)*	
Inputs	Recalls		*6.1.8 Recalls and Memory Modes*	
Inputs	Pedestrian Phasing	*5.1.1 Movement and Phase Numbering*		
Inputs	Walk Interval		*6.1.6 Pedestrian Intervals*	*7.5.1 Pedestrian Timing and Walk Modes*

	Signal Timing Parameter	Chapter 5 Reference	Chapter 6 Reference	Chapter 7 Reference
Inputs	Flashing Don't Walk Interval		*6.1.6 Pedestrian Intervals*	*7.5.1 Pedestrian Timing and Walk Modes*
	Dual Entry		*6.1.7 Dual Entry*	
	Inhibit Max			*7.4.3 Splits Guidance* *7.4.4 Force-Offs Guidance*
	Force-Offs			*7.3.4 Force-Offs* *7.4.4 Force-Offs Guidance*
	Offset Reference Point			*7.3.8 Offset Reference Point* *7.4.8 Offset Reference Point Guidance*
	Coordinated Phases			*7.3.1 Coordinated Phases* *7.4.1 Coordinated Phases Guidance*
	Yield Point			*7.3.6 Yield Point* *7.4.6 Yield Point Guidance*
	Detector Locations and Settings	*5.1.5 Detector Assignments*	*6.2 Detector Configurations*	*7.5.2 Actuating the Coordinated Phase*
Outputs	Cycle Length			*7.3.2 Cycle Length* *7.4.2 Cycle Length Guidance*
	Splits			*7.3.3 Splits* *7.4.3 Splits Guidance*
	Offsets			*7.3.9 Offsets* *7.4.9 Offsets Guidance*

5.4 REFERENCES

1. *Highway Capacity Manual 2010.* Transportation Research Board of the National Academies, Washington, D.C., 2010.
2. Chandler, B. E., M. C. Myers, J. E. Atkinson, T. E. Bryer, R. Retting, J. Smithline, J. T. P. Wojtkiewicz, G. B. Thomas, S. P. Venglar, S. Sunkari, B. J. Malone, and P. Izadpanah. *Signalized Intersections Informational Guide*, Second Edition. Report FHWA-SA-13-027, Federal Highway Administration, United States Department of Transportation, 2013.
3. Alexiadis, V., K. Jeannotte, and A. Chandra. *Traffic Analysis Toolbox Volume I: Traffic Analysis Tools Primer*. Report FHWA-HRT-04-038, Federal Highway Administration, United States Department of Transportation, 2004.

CHAPTER 6
INTERSECTION/UNCOORDINATED TIMING

CONTENTS

LIST OF EXHIBITS

CHAPTER 6. INTERSECTION/UNCOORDINATED TIMING

Chapter 6 provides guidance on basic signal timing parameters used at uncoordinated intersections (i.e., intersections running in "free" operation). Using the typical timing values in this chapter, a practitioner should be able to develop timing plans for a standard eight-phase intersection (without collecting extensive data or conducting extensive analysis). These timing parameters also apply when the controller is being coordinated with other intersections, but additional parameters, discussed in Chapter 7, must also be selected for coordinated operations.

6.1 BASIC SIGNAL TIMING PARAMETERS

The signal controller phase table contains signal timing parameter values associated with specific phases.

Several basic timing parameters are typically programmed for each phase at a signalized intersection (i.e., associated with the signal controller phase table), including

- Yellow change,
- Red clearance,
- Minimum green,
- Maximum green,
- Passage time (unit extension or gap time),
- Pedestrian intervals,
- Dual entry, and
- Recalls and memory modes.

"Interval" is a term that describes the time when a signal indication does not change (e.g., walk, flashing don't walk, or green interval).

This chapter includes descriptions of these parameters and their functions, as well as some direction for a practitioner to use as he or she chooses specific values. This guidance was developed assuming an eight-phase intersection with (1) full actuation for vehicles and pedestrians and (2) presence mode detection (discussed in detail in Section 6.2). Exhibit 6-1 provides a summary of the consequences associated with giving too much or too little time to a parameter, as well as the variables that should affect the ultimate parameter value that is chosen.

Exhibit 6-1 Basic Signal Timing Parameter Guidance

	Timing Parameter	Consequence for Too Little Time	Consequence for Too Much Time	Dependent On Variables Including:
Section **6.1.1**	Yellow Change	☐ May create a dilemma zone (Type I) ☐ May cause a higher frequency of red-light running	☐ May encourage disrespect by familiar drivers	☐ Driver perception-reaction time ☐ Vehicle deceleration rate ☐ Vehicle approach speed ☐ Approach grade
Section **6.1.2**	Red Clearance	☐ Potential conflict after phase begins	☐ Wasted time at the intersection	☐ Intersection width ☐ Vehicle length ☐ Vehicle approach speed

	Timing Parameter	Consequence for Too Little Time	Consequence for Too Much Time	Dependent On Variables Including:
Section 6.1.3	Minimum Green	□ May violate driver expectations (leading to a possible increase in rear-end crashes) □ May not accommodate pedestrian needs □ May not accommodate bicycle needs	□ Wasted time at the intersection	□ Driver expectancy □ Detector locations □ Number of queued vehicles □ Pedestrian intervals □ Bicycle speed and acceleration
Section 6.1.4	Maximum Green	□ Some vehicles may not be served because the phase capacity is inadequate for demand	□ Wasted time at the intersection (particularly if there is broken detection) □ Possible queuing on movements with long delays	□ Vehicle demand □ Intersection capacity
Section 6.1.5	Passage Time (Unit Extension or Gap Time)	□ Green may end prematurely before all vehicles have been served	□ Delays to other movements caused by extension of the phase	□ Detection design □ Detection mode □ Vehicle approach speed
Section 6.1.6	Walk	□ May not accommodate high volumes of pedestrians	□ Wasted time at the intersection	□ Pedestrian volumes □ Push button locations □ Pedestrian crossing distance □ Pedestrian walking speed
	Flashing Don't Walk (FDW)	□ May not accommodate the time needed for pedestrians to cross the street	□ Wasted time at the intersection	□ Pedestrian crossing distance □ Pedestrian walking speed

6.1.1 Yellow Change

The yellow change interval warns users that there is about to be a change in right-of-way assignment at the intersection.

6.1.1.1 Operating Environment Considerations for Yellow Change

A state's vehicle code fits into one of two broad categories—***permissive*** or ***restrictive*** yellow law—described below:

- ***Permissive Yellow Law:*** Vehicular traffic facing a steady circular yellow or yellow arrow signal is thereby warned that the related green movement is being terminated or that a red indication will be exhibited immediately thereafter. This rule comes from paragraph 11-202 of the *Uniform Vehicle Code* (*1*).
- ***Restrictive Yellow Law:*** These laws, which are less common, require stopping unless not safe to do so. An adequate yellow interval should allow vehicles to stop.

It is very desirable for drivers in a region to see a consistent application of the yellow change interval.

Because of the various interpretations of the yellow change interval, practitioners are encouraged to refer to local and regional statutes for guidance when determining the purpose of yellow change time.

6.1.1.2 Typical Values for Yellow Change

The *Manual on Uniform Traffic Control Devices* (*2*) requires that the duration of the yellow change interval be determined using engineering practices. These should be based on individual intersection conditions (i.e., approach speed). The yellow change interval should last approximately 3 to 6 seconds, with longer intervals being used on higher speed approaches.

The Institute of Transportation Engineers (ITE) offers Equation 6-1 for computing the yellow change interval (*3*). The equation calculates the time required for a driver to make a decision to come to a safe stop or proceed. This is the minimum time to eliminate a dilemma zone (Type I), which exists if the yellow is too short. Decision zones (also known as Type II dilemma zones or indecision zones), which are related to detection design, are discussed in detail in Chapter 4.

Equation 6-1

$$Y = t + \frac{1.47v}{2(a + 32.2g)}$$

where

Y = yellow change interval (seconds),

t = perception-reaction time to the onset of a yellow indication (seconds),

v = approach speed (miles per hour [mph]),

a = deceleration rate in response to the onset of a yellow indication (feet per second per second), and

g = grade, with uphill positive and downhill negative (percent grade/100) (feet/feet).

Although values will vary by user population and local conditions, a perception-reaction time (t) of 1.0 second and a deceleration rate (a) of 10 feet per second per second are often cited for use in Equation 6-1 (*4*, *5*) and have recently been recommended for use in *NCHRP 731* (*6*). These recommended values are different from those cited in highway geometric design policy documents because they are based on driver response to the yellow indication, which is an expected condition. They are not based on the longer reaction time necessary for an unexpected (or surprise) condition.

When applying Equation 6-1 to through movement phases, the speed used is generally either the 85th-percentile speed or the posted regulatory speed limit, depending on agency policy (*7*). When applying Equation 6-1 to left-turn phases, the speed can equal that of the adjacent through movement, but it can also be slower, as left-turning drivers inherently slow to a comfortable turning speed.

Exhibit 6-2 shows minimum yellow change interval results based on Equation 6-1 and the recommended values discussed above. These values for minimum yellow are long enough to prevent a dilemma zone, so drivers will not find themselves in an area where they can neither enter on yellow nor stop. This is the minimum yellow required to prevent a vehicle from entering on red when, in fact, the driver might not have had enough time to stop.

The values for the yellow change interval in Exhibit 6-2 are based on negligible approach grades. They should be increased by 0.1 second for every 1 percent of downgrade. Similarly, they should be decreased by 0.1 second for every 1 percent of upgrade. To illustrate, consider an approach with a 30 mph approach speed and 4-percent downgrade. The estimated yellow change period should be 3.6 seconds (= 3.2 + (0.1 × 4)).

Exhibit 6-2 Duration of Minimum Yellow Change Interval

Approach Speed (MPH)	Minimum Yellow Change[1] (Seconds)
25	3.0*
30	3.2
35	3.6
40	3.9
45	4.3
50	4.7
55	5.0
60	5.4

[1] Based on negligible approach grades. Adjustments are required for upgrades and downgrades.
* The MUTCD (*2*) recommends a minimum duration of 3 seconds for the yellow change interval.

6.1.2 Red Clearance

The red clearance interval is an optional signal timing parameter that provides a period at the end of the yellow change interval, during which the phase has a red signal display before the display of green for the following phase. The purpose of this interval is to allow time for vehicles that entered the intersection during the yellow change interval to reach an appropriate location prior to the next phase. Practitioners should understand that site-specific conditions, such as heavy vehicles and downgrades, may warrant special attention.

Some publications incorrectly refer to the red clearance interval as an all-red interval. Red clearance only applies to a single phase.

6.1.2.1 Operating Environment Considerations for Red Clearance

The use of a red clearance interval is optional, and there is no consensus on its application or duration. Recent research indicates that the use of a red clearance interval shows some reduction of red-light-running violations. In these studies, there was also a significant reduction in right-angle crashes after implementing a red clearance interval. However, other research suggests that this reduction may only be temporary. A comprehensive study of long-term effects for the Minnesota Department of Transportation (*8*) indicated short-term reductions in crash rates were achieved (approximately one year after the implementation), but long-term reductions were not observed, which implies that there may not be safety benefits associated with increased red clearance intervals.

A disadvantage of using the red clearance interval is that there is a reduction in available green time for other phases. At intersections where the timing for minor movements is restricted (e.g., under coordinated operations, which are explained in Chapter 7), the extra time for a red clearance interval comes from the remaining phases at the intersection. In cases where major movements are already at or near saturation, the reduction in capacity associated with providing red clearance intervals should be accounted for in an operational analysis.

Recent research associated with *NCHRP 731* (*6*) has confirmed that vehicles do not enter an intersection until 1 second after the start of green and recommends red clearance intervals as shown in Exhibit 6-3. (The width of the intersection is measured from the stop bar to the extension of the cross-street curb line or the outside edge of the farthest cross-street travel lane.) Some modern controllers can be configured to implement a red clearance extension when a vehicle is detected to be in an undesirable position at the onset of a conflicting green. This application is consistent with MUTCD guidance that allows lengthening of the red clearance interval.

Exhibit 6-3 Red Clearance Interval

Approach Speed (MPH)	Red Clearance[1] (Seconds)				
	Width of Intersection (Feet)				
	30	*50*	*70*	*90*	*110*
25	0.4	0.9	1.5	2.0	2.5
30	0.1	0.6	1.0	1.5	2.0
35	0.0	0.4	0.8	1.1	1.5
40	0.0	0.2	0.5	0.9	1.2
45	0.0	0.1	0.4	0.7	1.0
50	0.0	0.0	0.2	0.5	0.8
55	0.0	0.0	0.1	0.4	0.6
60	0.0	0.0	0.0	0.2	0.5

[1] Based on recent research reported in *NCHRP 731* (*6*), the calculated red clearance values have been reduced by 1 second.

6.1.3 Minimum Green

The minimum green parameter represents the least amount of time that a green signal indication will be displayed for a phase. Minimum green should be set to meet driver expectations, but its duration may also be based on considerations of queue length or pedestrian timing. High-speed rural locations, especially those with high truck volumes, may benefit from longer times than those at typical urban intersections.

A minimum green that is too long may result in increased delay and excessive queues at an intersection; one that is too short may violate driver expectations or, in some cases, pedestrian needs. (In this chapter, pedestrian detection and indications are assumed to be present. Pedestrian timing considerations are discussed in detail in Section 6.1.6.)

6.1.3.1 Operating Environment Considerations for Minimum Green

Vehicles that have longer start-up times (such as bicycles, trucks, and transit) generally require more time to clear queues at the beginning of the green interval. Longer minimum green times may be appropriate at locations with high numbers of these types of vehicles, depending on the type of detection design.

6.1.3.2 Typical Values for Minimum Green Based on Driver Expectancy

The recommended duration of minimum green needed to satisfy driver expectancy varies among practitioners. Some use 15 seconds or more of minimum green; other practitioners use as little as 2 seconds. If a minimum green parameter is set too low and violates driver expectancy, there is a risk of increased rear-end crashes. The values listed in Exhibit 6-4 are typical for the specified combinations of phase and facility type.

Exhibit 6-4 Typical Values for Minimum Green to Satisfy Driver Expectancy

Phase Type	Facility Type	Minimum Green (Seconds)
Through	Major Arterial (> 40 mph)	10 to 15
	Major Arterial (≤ 40 mph)	7 to 15
	Minor Arterial	4 to 10
	Collector, Local, or Driveway	2 to 10
Left Turn	Any	2 to 5

6.1.3.3 Typical Values for Minimum Green Based on Queue Clearance

As described in Chapter 5, start-up lost time is the time used by the first few vehicles in a queue to react to the green indication and accelerate.

In addition to driver expectancy, the duration of minimum green can also be influenced by the location of detectors. If stop bar detection is not present, a minimum green interval is needed to clear the vehicles queued between the stop bar and the nearest setback detector. If the minimum green is not long enough to clear the vehicles, a vehicle might get “stuck” between the setback detector and the stop bar. Without stop bar detection, the controller will not know that there is a vehicle waiting. Equation 6-2 (in combination with Equation 6-3) can be used to estimate the minimum green needed to satisfy queue clearance. Equation 6-2 assumes a start-up lost time of 3 seconds and that each subsequent vehicle requires 2 seconds to clear the intersection:

Equation 6-2

$$G_q = 3 + 2n$$

where

G_q = minimum green duration for queue clearance (seconds) and

n = number of vehicles between stop bar and nearest setback detector in one lane.

Equation 6-3

$$n = \frac{d}{L_v}$$

where

d = distance between the stop bar and the downstream edge of the nearest setback detector (feet) and

L_v = length of vehicle (feet), *set at 25 feet*.

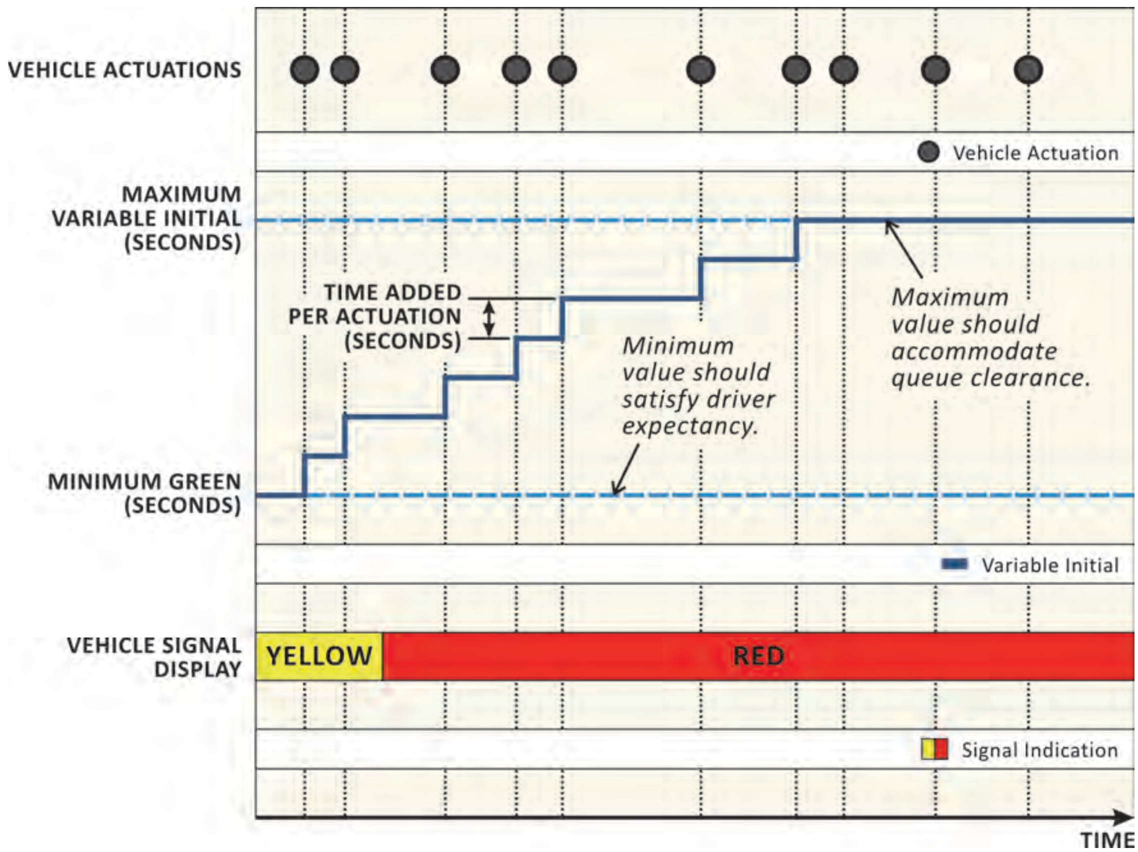
Modern controllers can adjust the amount of green time given to a phase based on queue length (when stop bar detection is not used) using a parameter called “variable initial.” The minimum green to satisfy driver expectancy (discussed in the previous section) is often used to define the lower limit for variable initial, and the maximum value is defined by the green time needed for queue clearance (provided it is higher than the minimum green required for driver expectancy). Variable initial will be increased incrementally based on the number of actuations (i.e., number of users being detected) during the yellow and red intervals until the upper limit is reached, but will never be less than the minimum green. Exhibit 6-5 provides some typical values for minimum green, maximum variable initial, and seconds added per actuation, while Exhibit 6-6 illustrates the concept. Note that the green time for queue clearance (calculated using Equation 6-2) is noted in Exhibit 6-5 as maximum variable initial.

Exhibit 6-5 Typical Variable Initial Timing Values for Setback Detection Queue Clearance (Without Stop Bar Detection)

Distance Between Stop Bar and Nearest Setback Detector (Feet)	Minimum Green (Seconds)	Maximum Variable Initial (Seconds)	Seconds Added per Actuation[1]
275	10	25	2.0
350	10	31	2.0
425	10	37	2.0
500	10	43	2.0

[1] Seconds added per actuation assumes approximately 2-second headways.

Exhibit 6-6 Variable Initial



6.1.3.4 Typical Values for Minimum Green Based on Bicycle Timing

At intersections with significant bicycle volumes and stop bar detection for bicycles, the minimum green time (along with the yellow change and red clearance intervals) should accommodate a bicycle crossing the intersection. A practitioner does not generally need to accommodate a bicycle solely within the minimum green time. Instead, bicycles should be accommodated within a minimum phase length (that includes a minimum green time, yellow change interval, and red clearance interval). Equation 6-4, provided by the California Department of Transportation's (Caltrans) *California Manual on Uniform Traffic Control Devices* (*9*), can be used to calculate the minimum amount of time required for a bicycle to clear an intersection.

$$Minimum\ Phase\ Length = G_{min} + Y + R_{clear} = 6\ sec + \frac{w + 6\ feet}{14.7\ feet/sec}$$

Equation 6-4

where

G_{min} = length of minimum green interval (seconds),

Y = length of yellow change interval (seconds),

R_{clear} = length of red clearance interval (seconds), and

w = distance from limit line to far side of last conflicting lane (feet).

A bicycle crossing speed of 10 miles per hour (14.7 feet per second) is assumed in the equation, and an effective start-up lost time of 6 seconds is used to represent the time lost to reacting to the green indication and then accelerating to full speed. Typical values for the minimum phase length used in a bicycle environment are provided in Exhibit 6-7.

Exhibit 6-7 Typical Values for Minimum Phase Length for Bicycle Timing

Bicycle Crossing Distance (Distance from Limit Line to Far Side of Last Conflicting Lane) (Feet)	Minimum Phase Length (Seconds)
40	9.1
50	9.8
60	10.5
70	11.2
80	11.9
90	12.5
100	13.2
110	13.9
120	14.6
130	15.3
140	15.9
150	16.6
160	17.3
170	18.0
180	18.7

6.1.4 Maximum Green

The maximum green parameter represents the maximum amount of time that a green signal indication can be displayed in the presence of a call on a conflicting phase. A maximum green that is too long may result in wasted time at an intersection, and movements that experience long delays as a result could experience queuing. If the maximum green value is too short, then the phase capacity may be inadequate for the traffic demand, and some vehicles will remain unserved at the end of the green interval. As shown in Exhibit 6-8, the maximum green timer begins timing upon the presence of a conflicting call. If there is demand on the phase that is currently timing and there are no conflicting calls, the maximum green timer will be reset to its maximum value until an opposing call occurs.

"Call" is a term used to describe the presence of vehicles or pedestrians at detectors.

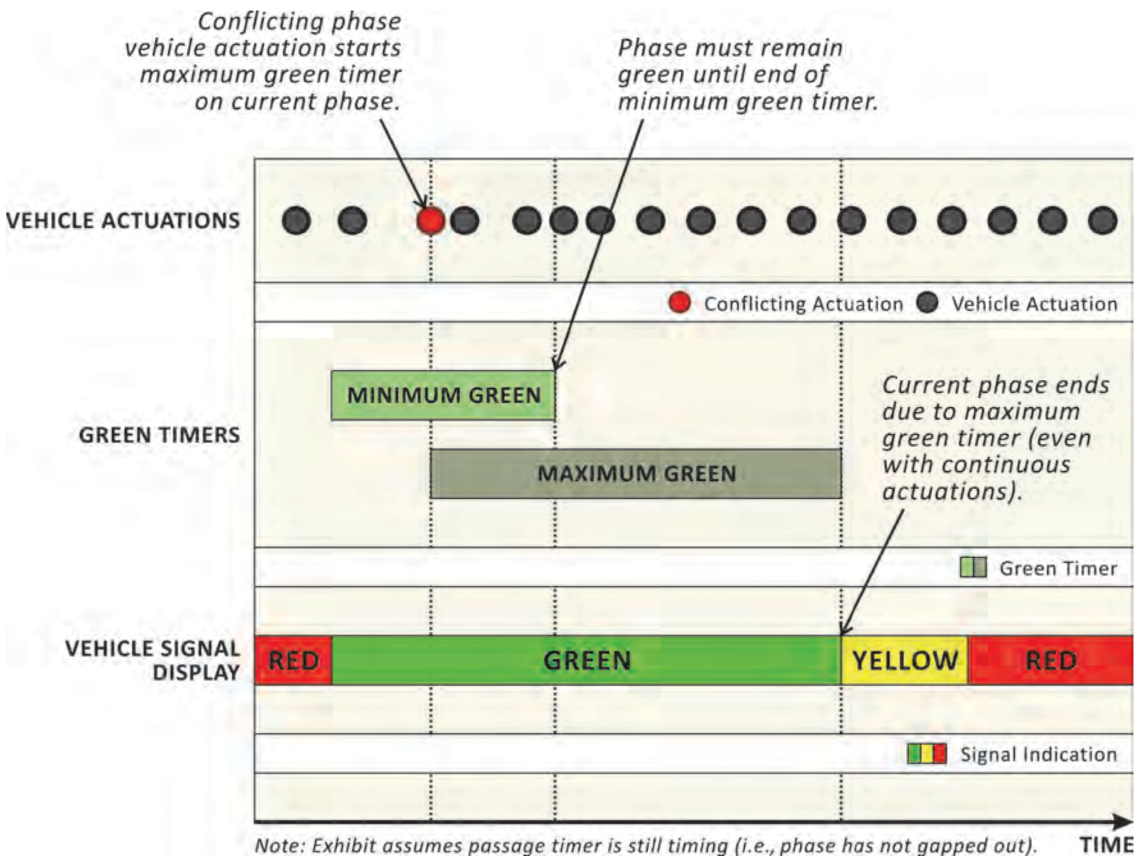
Maximum green is used to limit the delay to any movement at an intersection and keeps the cycle length to a desired maximum amount. It also guards against long green times due to continuous demand or broken detectors. The normal failure mode of a detector is to place a continuous call for service, so a failed detector on a phase will cause that phase's maximum green to time every cycle. Therefore, setting a proper

When a phase terminates at its maximum green time (instead of during a gap in traffic), the phase is considered to have "maxed out."

maximum green duration could potentially minimize the impact to traffic in case of detector failure.

Ideally, the maximum green will not be reached because the detection system will find a gap to end the phase (discussed in detail in Section 6.1.5). However, if there are continuous calls for service on the current phase and a call on one or more conflicting phases, the maximum green parameter will eventually terminate the current phase.

Exhibit 6-8 Maximum Green



6.1.4.1 Operating Environment Considerations for Maximum Green

The maximum green value is typically developed based on traffic volumes and does not generally change with different operating environment characteristics. One exception is on residential minor streets where long maximum greens might encourage cut-through traffic.

6.1.4.2 Typical Values for Maximum Green

The maximum green value should exceed the green duration needed to serve the typical maximum queue and, thereby, allow the phase to accommodate cycle-to-cycle peaks in demand. A properly timed maximum green duration will commonly result in frequent phase termination by gap out (where a gap in traffic, indicating inefficient flow, results in the end of a phase) during low-to-moderate volumes and by occasional max out during peak periods.

Exhibit 6-9 provides some typical ranges for maximum green based on various facility types. Maximum green values shown in Exhibit 6-9 should be used as a starting point and adjusted based on field conditions. Good practice is to vary maximum green duration by time of day because traffic volumes could fluctuate quite significantly between peak and off-peak hours. (Time-of-day plans are explained in detail in Section 6.3.)

Phase Type	Facility Type	Maximum Green (Seconds)
Through	Major Arterial (> 40 mph)	50 to 70
	Major Arterial (≤ 40 mph)	40 to 60
	Minor Arterial	30 to 50
	Collector, Local, or Driveway	20 to 40
Left Turn	Any	15 to 30

Exhibit 6-9 Typical Values for Maximum Green Based on Facility Type

6.1.5 Passage Time (Unit Extension or Gap Time)

Passage time (also called unit extension or gap time) is a parameter that can be used to terminate the current phase when a gap in traffic is identified. A gap indicates that traffic is no longer operating at efficient flow rates and is often associated with headways greater than 2 to 3 seconds in each lane. Vehicle calls will extend green time on the current phase until a gap in detector occupancy is greater than the assigned passage time. If volumes are low enough, passage time allows a phase to end prior to its maximum green time. If the passage time is too short, the green may end prematurely—before vehicles have been adequately served. If the passage interval is set too long, there will be delays to other movements caused by unnecessary extension of the phase.

"Headway" refers to the time between two successive vehicles as they pass a point on the roadway, measured from a common vehicle feature (e.g., front axle or front bumper).

Passage time essentially operates through a timer that starts to time down from the instant detector actuation is removed (i.e., a user leaves the detector). In a modern controller, the passage timer is constantly running and expires when it times out. A subsequent actuation (i.e., another user crossing the detector) will reset the passage timer if it has not yet expired. The green interval can be extended up to the maximum green time while the passage timer is still timing. The current phase will only gap out (i.e., terminate before the maximum green time) if all of the following conditions are met:

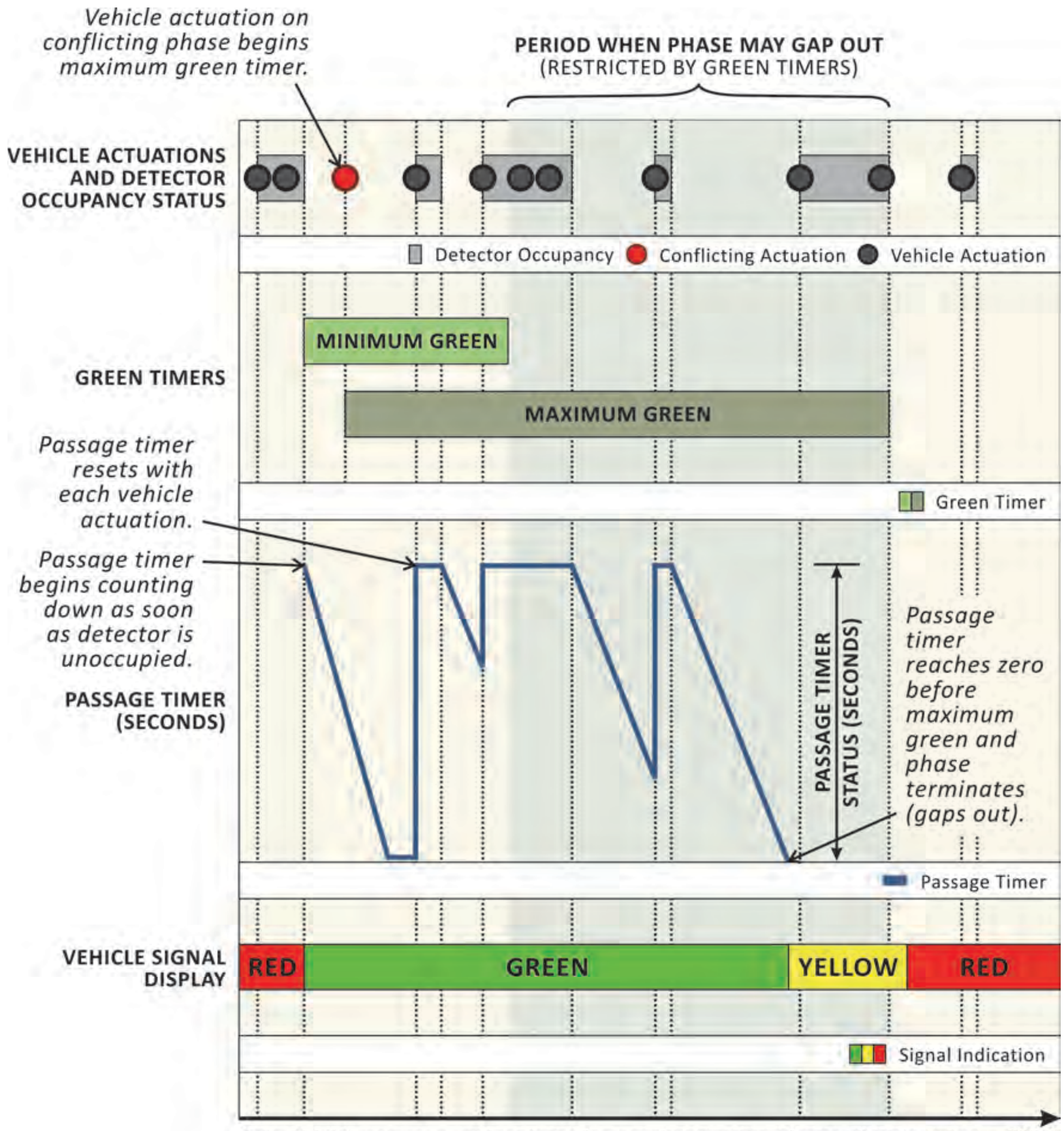
When the passage timer expires (times out), it is commonly referred to as a "gap out."

1. The minimum green timer has expired.
2. A call is waiting for service on a conflicting phase.
3. The passage timer has expired.

Exhibit 6-10 illustrates vehicles moving over a typical setback detector and shows how that affects the passage timer and the eventual termination of the phase through a gap out.

The mode of the detector (pulse or presence) is extremely important for passage time. Pulse mode provides a single pulse (only resetting after a vehicle leaves the detector), so the passage timer can time out with a car sitting over the detector. For this reason, pulse mode is not typically used in traffic signal control, except when system detectors are placed only to count vehicles (where queuing does not take place). If system detectors are used to measure occupancy, they must be in presence mode and should be located beyond normal intersection queues.

Exhibit 6-10 Passage Time



Note: Exhibit assumes no other timers (e.g., pedestrian timers) are still timing when phase gaps out.

When using presence mode, the speed of the vehicles crossing the detectors and the size of the detectors are important considerations for determining passage time. Large area detection (60 to 80 feet) allows passage times to be reduced to near zero, providing more efficient stop bar detection without fear of gap out, as even slow-moving vehicles typically do not leave a 60-foot gap. When stop bar detectors have passage time, there is effective "lost time" when the last vehicle departs the detection zone, as the signal does not turn yellow until after the vehicle has left the stop bar.

In a typical dual-ring phasing sequence with leading left-turn phases, both passage timers for the two concurrent phases (typically two through phases) preceding the barrier need to time out before phase termination (by gap out) can occur. If the "simultaneous gap out" feature is used, an expired passage timer (that has reached zero) can be reset with subsequent actuation prior to phase termination if the phase in the other ring has not gapped out. However, simultaneous gap out is not recommended in medium- to high-volume environments because it does not provide efficient operations. Note that an expired passage timer can be reset before the end of the minimum green interval regardless of the simultaneous gap out setting.

The relationship between passage time and flow rate is illustrated in Exhibit 6-11. Initially, flow rate increases at the beginning of green to a maximum rate (following start-up lost time). In this example, the maximum flow rate is shown at 1900 vehicles per hour. This example is meant to show how a change in passage time will change the point at which the phase ends. For example, a passage time of 2 seconds translates into a flow rate of approximately 1800 vehicles per hour. Given a passage time of 2 seconds, the passage timer will not expire until 2 seconds after the flow rate drops below 1800 vehicles per hour. So if this drop in flow rate (below 1800 vehicles per hour) occurs at 15 seconds after the beginning of green, the traffic signal will turn yellow at 17 seconds. If the passage time is 3 seconds, then the passage timer will expire at 17 seconds with this flow rate profile, and the traffic signal will turn yellow at 20 seconds. Note that the flow rate values are for example purposes; intersections will have different flow rate profiles. While the flow rates associated with each passage time will remain constant, the point at which the flow rate drops below those constant values will depend on local conditions.

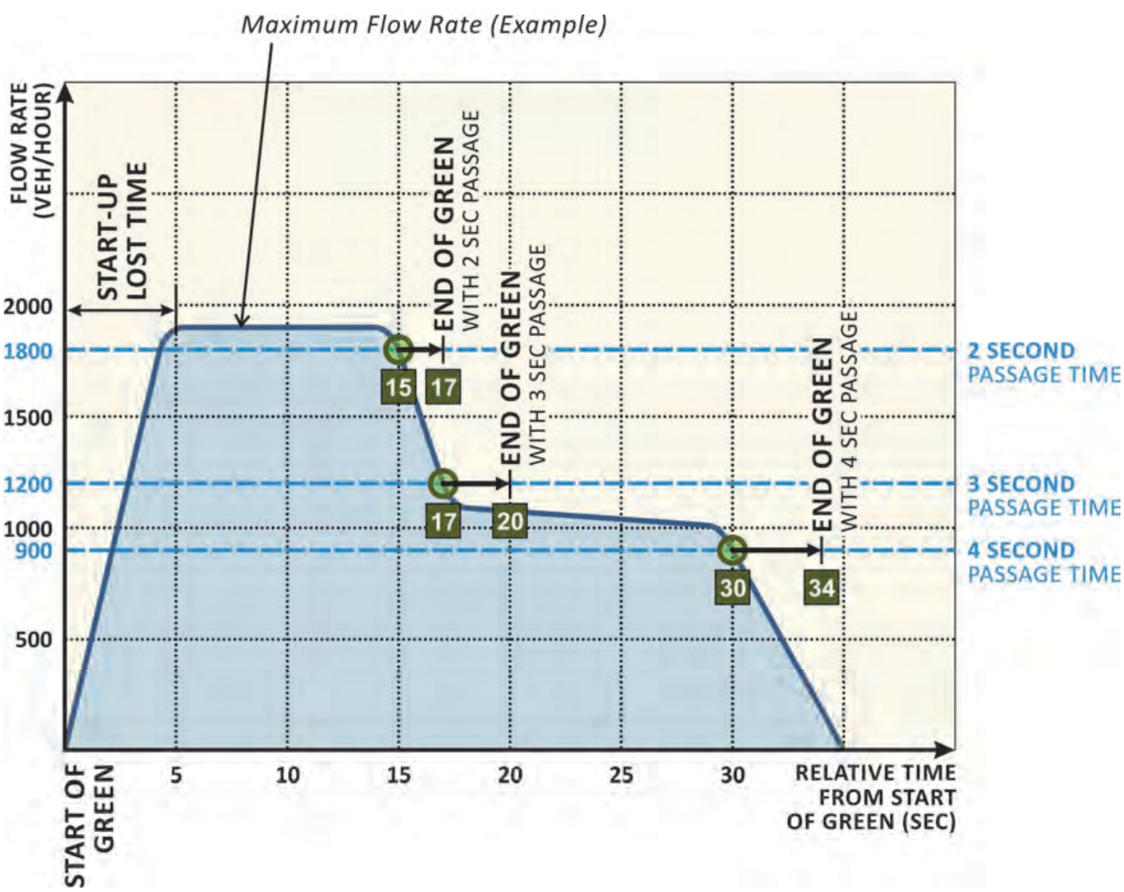


Exhibit 6-11 Passage Time and Flow Rate Relationship

The critical point to understand is that the passage timer is directly related to efficiency at a signalized intersection. The challenge is getting as close to an optimum value of passage time as practical without short-timing a phase. (See Section 6.1.5.2 for information on the gap reduction feature.) If traffic is light, giving a few extra seconds to less than efficient operations (i.e., flows below the maximum rate) may have minimal impact on other traffic. However, if the intersection is congested on any approach, a few extra seconds of inefficient flow may be detrimental to other movements. As the passage time increases, the amount of inefficient flow increases because the traffic signal will remain green for the length of selected passage time after the last vehicle is detected at the associated flow rate. Even if there are no vehicles after the last vehicle is detected at the measured flow rate, the traffic signal will remain green for the amount of passage time.

6.1.5.1 Operating Environment Considerations for Passage Time

Pedestrian actuations do not affect the passage time; pedestrian intervals will time regardless of whether the passage timer is timing or has expired. Bicycles are typically accommodated through an appropriate minimum green interval (see Exhibit 6-7) and are also not a factor when determining passage time. Under certain truck operating environment characteristics, increasing passage time to better accommodate trucks may be desirable depending on detection design. Considerations for trucks and transit can also be addressed with preferential treatment (discussed in Chapter 10).

6.1.5.2 Typical Values for Passage Time

The appropriate passage time used for a particular signal phase depends on many factors, including detection design (e.g., type and number of detection zones per lane, as well as location and size of detection zones), detection mode (presence recommended), and approach speed. Ideally, detection is designed and passage time selected to ensure that the system provides efficient queue service and safe phase termination on higher speed approaches. It is important to understand that the passage timer associated with a phase is based on a single detector in a single lane.

Exhibit 6-12 provides typical passage time values to sustain a flow rate of 1200 vehicles per hour (3-second headways) for a range of speeds and detection zone lengths. These values assume a single detector in a single lane (i.e., either a setback detector or a stop bar detector). Stop bar detectors are typically large area (e.g., 6 feet by 60 feet), and setback detectors are typically small area (e.g., 6 feet by 6 feet). If both setback detection and stop bar detection are available, the values in the table assume that the passage timer has been transferred exclusively to the setback detector. “Stop bar detector disconnect” is a feature in most modern controllers (and has a variety of different names). Setback detectors typically have passage times between 2 and 3 seconds if a single detector is used in each lane.

If multiple lanes are associated with a phase and the detector inputs are wired separately, passage time can be timed on the detector function of modern controllers. There are generally two locations within modern controllers where passage time can be assigned—the detector function and the phase table. Note that if passage time is placed on the detector function, the passage time in the phase table should be zero so as not to have two or more timers interacting. The practitioner should reference specific controller manuals for detailed information.

Exhibit 6-12 Typical Values for Passage Time

Detection Zone Length (Feet)	Passage Time (with a Headway of 3 Seconds) (Seconds)						
	Posted Speed (MPH)						
	25	*30*	*35*	*40*	*45*	*50*	*55*
6	2.3	2.4	2.5	2.6	2.6	2.6	2.7
20	1.9	2.1	2.2	2.3	2.4	2.5	2.5
40	1.4	1.6	1.8	2.0	2.1	2.2	2.3
60	0.8	1.2	1.4	1.6	1.8	1.9	2.0
80	0.3	0.7	1.1	1.3	1.5	1.6	1.8

As shown in Exhibit 6-12, the longer the detection zone, the sh orter the required passage time. Shorter detection zones increase the probability that a phase will gap out due to sluggish traffic. Slow-moving vehicles (such as trucks) may need a longer passage time in order to reach a short detection zone, reset the passage timer, and continue through the intersection before the phase times out. Zones that are 60 feet or longer will have virtually no probability of gapping out even with little or no passage time. Using longer detection zones and shorter passage times results in minimal "lost time" once the last vehicle departs the stop bar.

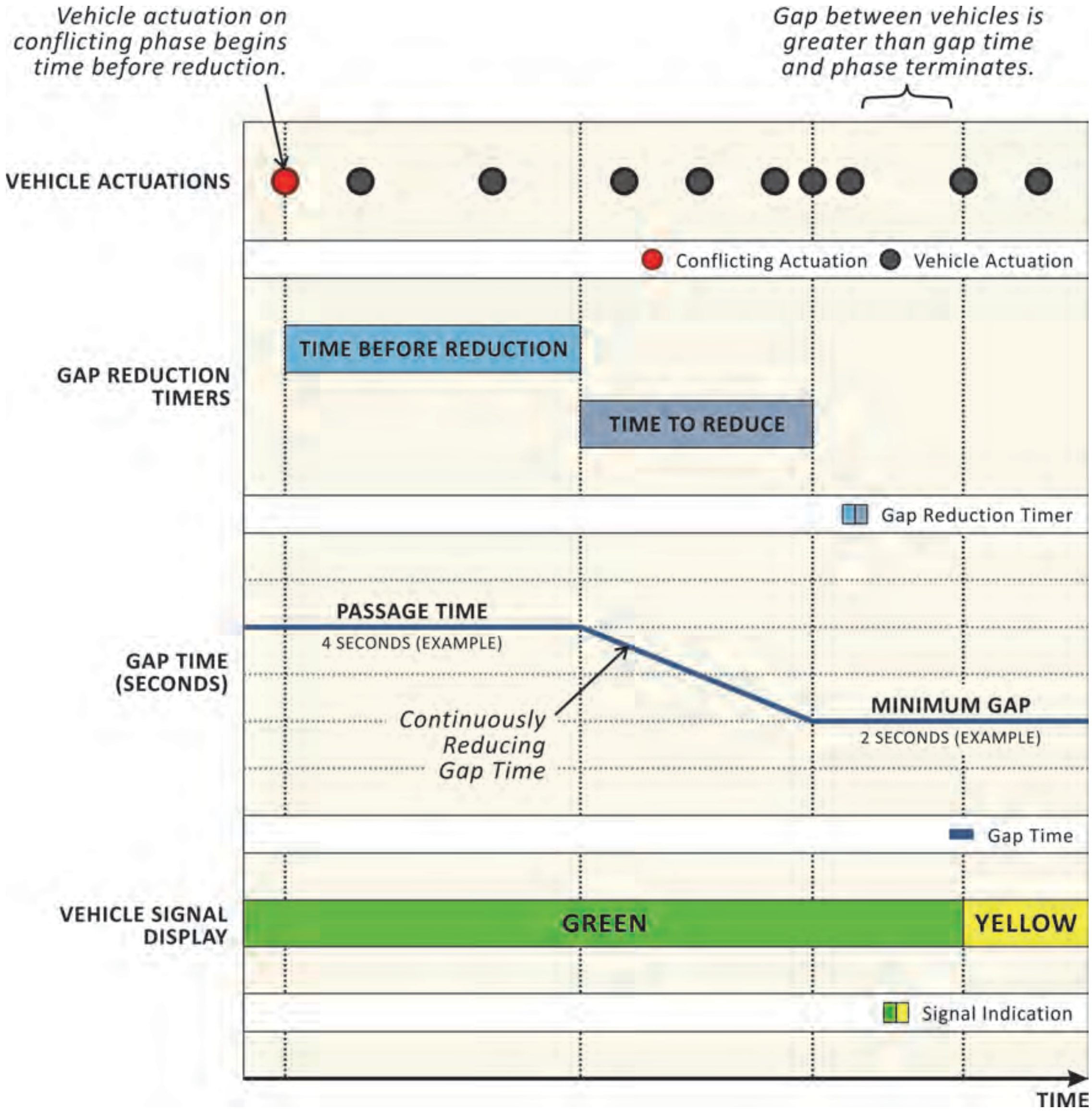
Values provided in Exhibit 6-12 can be used as a starting point, and if adjustments are desired, the following guiding principles should be followed (*10*):

1. ***Ensure queue clearance.*** The passage time should not be so small that the resulting headway causes the phase to have frequent premature gap outs unintentionally (i.e., a gap out that occurs before the queue is fully served). A premature gap out will leave a portion of the stopped queue unserved and, thereby, lead to increased delays and possible queue spillback.
2. ***Reduce max out frequency.*** The passage time should not be so large that the resulting operation causes the phase to have frequent max outs. A long passage time would allow even light traffic volumes to extend the green to max out. Users who are waiting in higher volume, conflicting phases may be delayed.

In certain high-volume environments, gap reduction features may be used for more efficient gap outs, especially on multi-lane approaches. With a gap reduction feature, a higher passage time is used initially to prevent premature gap outs when vehicles are slowly clearing the intersection. After a specified time ("time before reduction"), the passage time is reduced to a minimum gap value using a gradual reduction over a specified time ("time to reduce") as vehicular flow decreases (shown in Exhibit 6-13).

The time before reduction parameter establishes the time that is allowed to elapse after the arrival of a conflicting call and before the passage time begins to reduce. Once the time before reduction period expires, the passage time is reduced in a linear manner until the time to reduce period expires. Reducing the passage time to a minimum gap value allows the phase to gap out more easily as the phase progresses.

Exhibit 6-13
Relationship between Passage Time, Minimum Gap, Time before Reduction, and Time to Reduce



The gap reduction features can also be demonstrated using the flow rate concept introduced in Exhibit 6-11. The point at which the phase ends will depend on the flow rate profile, the gap reduction features, and conflicting traffic. Exhibit 6-14 illustrates how the use of a gap reduction feature (in this case, using a 3-second passage time and a 2-second minimum gap time) results in a phase termination between what could have been achieved with either a 2-second or 3-second passage time value.

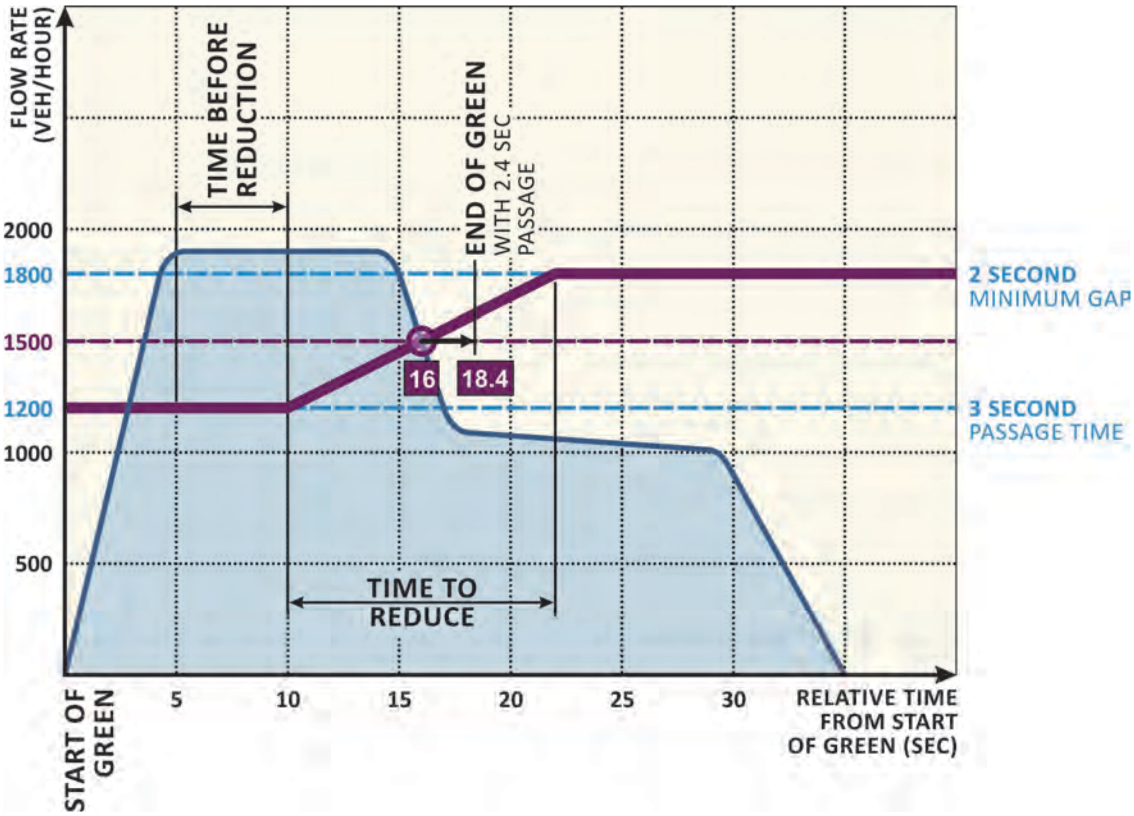


Exhibit 6-14 Gap Reduction Features and Flow Rate Relationship

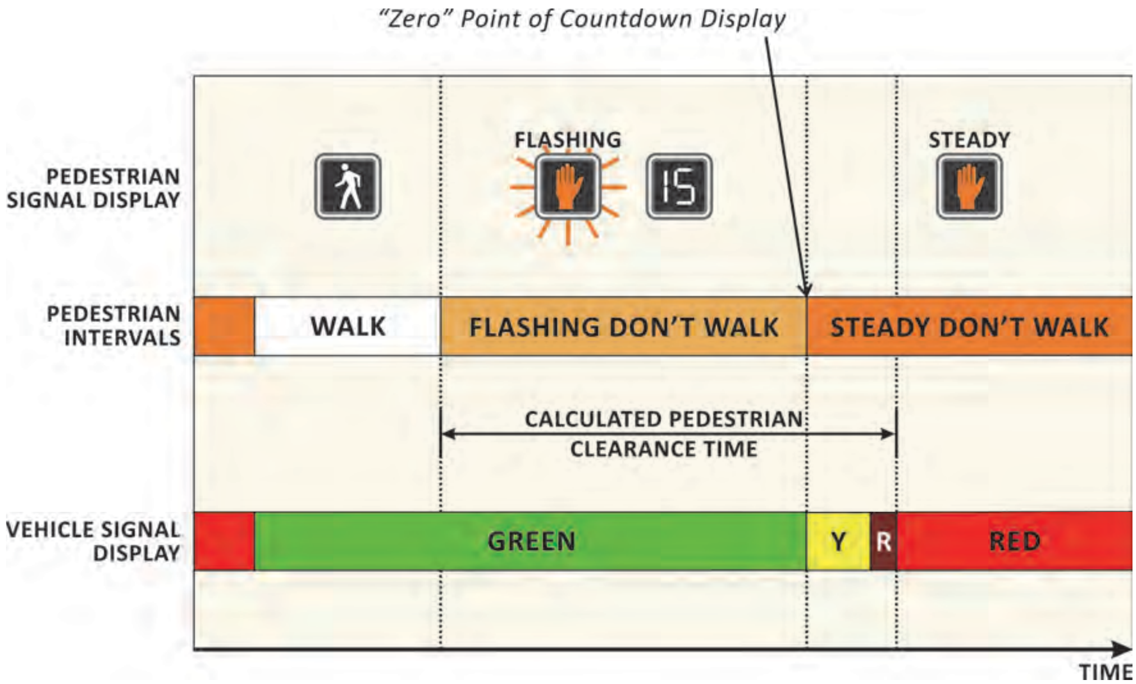
6.1.6 Pedestrian Intervals

The pedestrian phase consists of three intervals: walk, flashing don't walk (FDW), and steady don't walk (as shown in Exhibit 6-15). The walk interval typically begins at the start of the concurrent vehicular green interval and is used so that pedestrians can react to the start of the phase and move into the crosswalk. The FDW interval follows the walk interval and informs pedestrians that the phase is ending. Typically the FDW ends at the beginning of yellow, which is necessary for countdown pedestrian signals. (Note that ending the FDW at the beginning of yellow automatically satisfies the 3-second buffer interval required by the MUTCD [*2*].)

While the FDW interval may be called the pedestrian clearance interval in some controllers, this is not necessarily accurate. The pedestrian clearance interval is the time required to cross the street, and the FDW interval is typically the pedestrian clearance interval reduced by the yellow change and red clearance intervals. While some agencies may choose longer pedestrian clearance times that do not include yellow and red clearance, it is not required. The steady don't walk interval follows the FDW interval. The steady don't walk time is not a programmable parameter in the controller. The duration of the steady don't walk interval is simply the length of the phase minus the walk and FDW intervals. However, per the MUTCD, it must be displayed for at least 3 seconds before the release of any conflicting vehicular movements (*2*).

The pedestrian walk and FDW intervals typically time concurrently with vehicle intervals. The minimum duration of the walk is programmed in the controller; the actual walk interval may be longer depending on many other controller settings, which are discussed in Chapter 7. The practitioner must consider the influence of the pedestrian intervals on the other users at the intersection. If the pedestrian intervals require more time than required for vehicles or permitted by the maximum green timer, the vehicle phase will continue to time until the pedestrian FDW interval finishes timing.

Exhibit 6-15
Pedestrian Intervals



6.1.6.1 Operating Environment Considerations for Pedestrian Intervals

The length of pedestrian intervals may be increased based on pedestrian volumes, pedestrian characteristics (e.g., elderly pedestrians or school children) and facility characteristics (e.g., location of pedestrian push buttons and pedestrian crossing distances). A longer walk interval may be considered with high pedestrian volumes, and a longer FDW interval may be used near locations with users requiring extra time.

Leading pedestrian intervals may be considered with high pedestrian volumes or with certain pedestrian safety issues (such as high volumes of turning vehicles). Leading pedestrian intervals are a treatment that allows pedestrians to establish their presence in a crosswalk, which reduces interference from turning vehicles. More information is provided in Section 6.1.6.6.

6.1.6.2 Typical Values for Pedestrian Walk Interval

The length of the walk interval may be established in local agency policy.

The walk interval should provide pedestrians with adequate time to perceive the walk indication and depart the curb before the FDW interval begins. In other words, it should be long enough to allow a pedestrian or multiple pedestrians at high-pedestrian-volume locations to enter the crosswalk. MUTCD (*2*) guidance states that the walk interval should be at least 7 seconds in length. In areas with high pedestrian volumes (such as school zones, central business districts (CBDs), and sports and event venues), longer walk intervals should be considered. Typical values for the pedestrian walk interval are provided in Exhibit 6-16. These values are based on information from the MUTCD and *Traffic Control Devices Handbook* (*2*, *5*).

Conditions	Walk Interval (Seconds)
High-pedestrian-volume area (e.g., school, CBD, or sports and event venue)	10 to 15
Typical pedestrian volume and longer cycle length	7 to 10
Typical pedestrian volume and shorter cycle length	7
Negligible pedestrian volume and otherwise long cycle length	4

Exhibit 6-16 Typical Values for Pedestrian Walk Interval

6.1.6.3 Typical Values for Pedestrian Clearance

Pedestrian clearance is the time required for a pedestrian to complete crossing the street, assuming the walk interval has expired and the FDW interval has just begun. One special-use option is using pedestrian detection to automatically adjust the pedestrian clearance time based on actual pedestrian walking speeds or clearance of the crosswalk, but the more traditional process is to calculate the required pedestrian clearance using a predetermined walking speed.

The MUTCD (*2*) states that pedestrian clearance times should be sufficient to allow a pedestrian walking at a speed of 3.5 feet per second to cross from the curb or shoulder to at least (a) the far side of the traveled way or (b) to a median of sufficient width for pedestrians to wait (*2*). Where slower pedestrians routinely use a crossing, a walking speed less than 3.5 feet per second may be considered when determining the pedestrian clearance time. Alternatively, the MUTCD (*2*) outlines an option to use a walking speed of up to 4 feet per second to evaluate whether the pedestrian clearance time is sufficient. This option may be applied at locations where an extended push button press function has been installed to provide slower pedestrians an opportunity to request and receive a longer pedestrian clearance time. An example of the extended push button treatment signage is shown in Exhibit 6-17.

Exhibit 6-17 Example of Extended Push Button Press Function Signage

Using an established walking speed, pedestrian clearance time can be calculated using Equation 6-5. To calculate the FDW time that is routinely programmed in the signal controller phase table, the practitioner should subtract the yellow change and red clearance interval times from the calculated pedestrian clearance time.

$$PCT = \frac{D_c}{v_p}$$

Equation 6-5

where

PCT = pedestrian clearance time (seconds),

D_c = pedestrian crossing distance (feet), and

v_p = pedestrian walking speed (feet per second).

Typical values for pedestrian clearance time (calculated using Equation 6-5), assuming a walking speed of 3.5 feet per second, for various pedestrian crossing distances are shown in Exhibit 6-18.

Exhibit 6-18
Calculated Values for Pedestrian Clearance Time (Based on a Walking Speed of 3.5 Feet per Second)

Pedestrian Crossing Distance (Feet)	Calculated Pedestrian Clearance Time (Seconds)
40	11
60	17
80	23
100	29

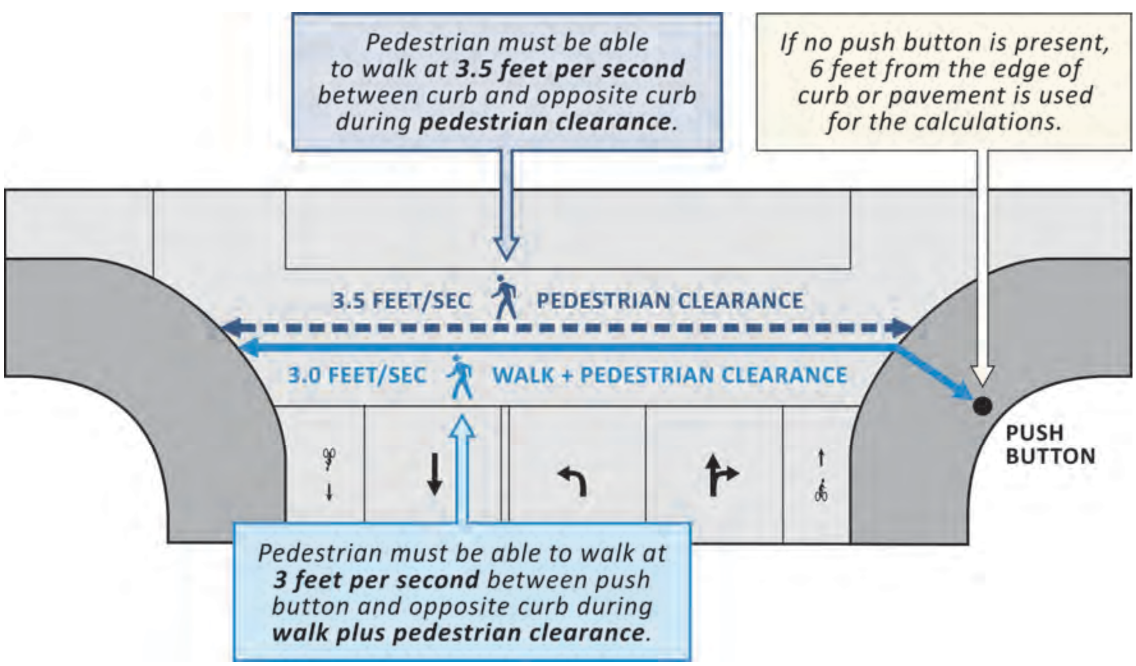
6.1.6.4 Typical Values for Pedestrian Flashing Don't Walk Interval

The pedestrian FDW interval (incorrectly called pedestrian clearance in some controller phase tables) can be determined by reducing the calculated intersection pedestrian clearance time by the yellow change and red clearance times. The FDW time displayed during the yellow change and red clearance intervals is essentially the "buffer" time before a conflicting movement. Some agencies may prefer a more conservative approach, and will not reduce the calculated pedestrian clearance time by the yellow change and red clearance times. A practitioner should verify standard practice through a review of jurisdiction policies.

6.1.6.5 Additional Guidance on Pedestrian Intervals—4-Second Walk

While pedestrian walk intervals of 7 seconds are usually appropriate in most situations, walk intervals as short as 4 seconds may be used (*2*) if pedestrian volumes and characteristics do not necessitate the full 7-second walk interval. MUTCD (*2*) guidance states that the total of the 4-second walk and pedestrian clearance (calculated at 3.5 feet per second) should be enough time for a pedestrian to travel from the pedestrian detector (or if no detector is present, a location 6 feet from the edge of curb or pavement) at the beginning of the walk signal indication to the far side of the traveled way or the median, while traveling in the crosswalk at a walking speed of 3 feet per second (see Exhibit 6-19). Any additional time that is required to satisfy the condition of the 3-feet-per-second guidance should be added to the walk interval (*2*).

Exhibit 6-19
Pedestrian Interval Requirements Based on Walking Speed



6.1.6.6 Additional Guidance on Pedestrian Intervals—Leading Pedestrian Interval

A leading pedestrian interval allows the walk indication for a pedestrian phase to be displayed prior to the associated vehicle phase. This treatment allows a pedestrian to establish right-of-way in an intersection and can also aid in pedestrian visibility for drivers, bicyclists, and other system users. In the case of special pedestrian treatments, the MUTCD (*2*) states that if a leading pedestrian interval is used, it should be at least 3 seconds in duration, and consideration should be given to prohibiting turns across the crosswalk during this time. Exhibit 6-20 uses a ring-and-barrier diagram to illustrate a phasing sequence with leading pedestrian intervals on Phases 2 and 6.

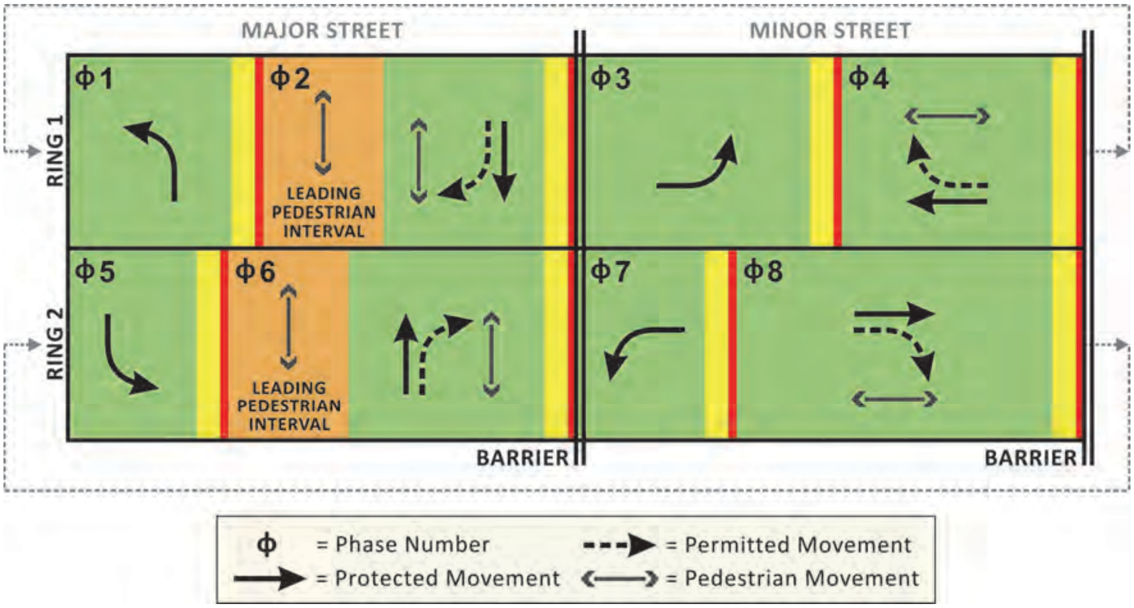


Exhibit 6-20 Leading Pedestrian Intervals

6.1.7 Dual Entry

The dual (double) entry parameter is used to call vehicle phases that can time concurrently, even if only one of the phases is receiving an active call. For example, if dual entry is active for Phases 4 and 8, and Phase 4 receives a call but no call is placed on Phase 8, Phase 8 will still be displayed along with Phase 4. The most common use of dual entry is to activate the parameter for compatible through movements. If the dual entry parameter is not selected, a vehicle call on a phase will only result in the timing of that phase in the absence of a call on the compatible phase.

6.1.8 Recalls and Memory Modes

The recall parameter causes the controller to place a call automatically for a specified phase regardless of the presence of any detector-actuated calls. There are four types of recalls: minimum recall (also known as vehicle recall), maximum recall, soft recall, and pedestrian recall. If no recalls or serviceable calls exist, the controller may be configured to rest in red or the last green. Descriptions are the following:

- ***Minimum Recall:*** The minimum recall parameter causes the controller to place a call for vehicle service on a phase in order to serve at least its minimum green duration.

Recall and memory modes are controller settings that place calls on a particular phase so that the phase will be served either automatically or based on past vehicle actuations, respectively.

- ***Maximum Recall:*** The maximum recall parameter causes the controller to place a continuous call for vehicle service on a phase in order to run its maximum green duration every cycle.
- ***Soft Recall:*** The soft recall parameter causes the controller to place a call for vehicle service on a phase *in the absence* of a serviceable conflicting call.
- ***Pedestrian Recall:*** The pedestrian recall parameter causes the controller to place a continuous call for pedestrian service on a phase, resulting in the controller timing its walk and FDW intervals every cycle.
- ***Red Rest:*** All phases may rest in red state when no serviceable calls exist in the ring and none of the above recall modes are used.
- ***Green Rest:*** If there is no serviceable call and no recall or red rest, the controller will rest in the last green.

Memory modes refer to the controller's ability to "remember" (i.e., retain) a detector actuation for a particular phase when an actuation is received during the red interval (and optionally, the yellow interval). One of two modes can be used—non-locking or locking. Memory mode is typically set on a per-phase basis, although some traffic signal controllers may have memory mode settings available for each detector channel input. The non-locking mode is typically the default mode, but regardless of the chosen detector mode, *all* actuations received during the *green* interval are treated as non-locking by the controller. Descriptions of the modes are the following:

- ***Non-Locking Mode:*** In the non-locking mode, an actuation received from a detector is not retained by the controller after the actuation is dropped by the detection unit. The controller recognizes the actuation only during the time that it is held present by the detection unit.
- ***Locking Mode:*** In the locking mode, the first actuation received by the controller on a specified channel during the red interval (and optionally, yellow interval) is used by the controller to trigger a continuous call for service. This call is retained until the assigned phase is served, regardless of whether any vehicles are waiting to be served.

6.1.8.1 Operating Environment Considerations for Recalls and Memory Modes

The practitioner should determine the recall setting (or allow red rest) based on the operating environment. For example, arterials are likely best served by soft recall if detection exists on all phases. Pedestrian recall, on the other hand, may be used at locations and/or times with high pedestrian volumes.

6.1.8.2 Typical Settings for Recalls and Memory Modes

Certain recall and memory mode settings may be more appropriate than others depending on the detection design and intended signal operations. Exhibit 6-21 provides a summary of typical detection designs and considerations for various recall and memory mode settings. This guidance assumes a single detector per lane. Intersections with more complicated detector designs would have different settings and need to be evaluated on a case-by-case basis.

Exhibit 6-21 Recall and Memory Mode Considerations

	Settings	Typical Detection Designs	Reasons for Use
Recalls	No Recalls	Stop bar detectors or setback detectors with locking mode	Relatively light traffic demand, and phases need not be served every cycle.
	Minimum Recall (Vehicle Recall)	Setback detectors with non-locking mode	Phases are expected to be served every cycle (e.g., on major street through movements) with green time to be extended based on vehicle actuations.
	Maximum Recall	No detectors	Fixed-time operation is desired, or gapping out is not desired. Maximum recall is typically not used at fully-actuated intersections with good detection systems and proper passage time settings.
	Soft Recall	Setback detectors with locking mode	Use on major street through movements to ensure these movements will dwell in green in the absence of calls on conflicting phases and allow major street through movements to be skipped in the absence of major street phase calls when calls exist on other phases.
	Pedestrian Recall	No pedestrian detectors	Pedestrian demand is expected to be high, and pedestrian phases are expected to be served every cycle.
	Red Rest	Stop bar detectors or setback detectors with locking mode	Typically at intersections with no definable major or minor streets.
Memory Modes	Non-Locking Mode	Stop bar detectors	Allows permitted movements (e.g., right-turn-on-red) to be completed without invoking a phase change.
	Locking Mode	Setback detectors with no stop bar detectors	Use on major street through movements associated with a low percentage of turning vehicles when recalls are not used.

Assuming Phases 2 and 6 are assigned to the major street through movements, Exhibit 6-22 shows the typical settings on recalls and memory modes at a fully-actuated intersection. At fully-actuated intersections where there are not definable major or minor streets, some agencies may prefer to dwell in the last phase, dwell in the next through phase, or dwell in red. This allows the intersection to dwell in the most appropriate state in the absence of any vehicle or pedestrian calls.

Exhibit 6-22 Typical Settings on Recalls and Memory Modes at a Fully-Actuated Intersection

	Settings	On for Phases	Off for Phases
Recalls	Minimum Recall (Vehicle Recall)	None	All Phases
	Maximum Recall	None	All Phases
	Soft Recall	2 and 6	All Phases (Except 2 and 6)
	Pedestrian Recall	None	All Phases
Memory Modes	Locking Mode[1]	None	All Phases

[1] Common names in traffic signal controllers include “Memory,” “Red Lock,” etc. “Yellow Lock” provides similar operations, but also allows actuations received during the yellow interval to be retained.

6.2 DETECTOR CONFIGURATIONS

This section explains the following detector parameters commonly used at signalized intersections:

- Delay,
- Extend/carry-over/stretch time, and
- Detector switching.

These detector parameters are only applicable to vehicle detectors (not pedestrian detectors).

6.2.1 Delay

Delay is used to temporarily disable detector output for a phase, essentially preventing vehicle actuations from being recognized right away by the signal controller. An actuation is not made available unless the delay timer has expired and the detection zone is still occupied. Delay can either be programmed on the detector or in the controller. However, the signal timing practitioner should be careful not to program delay into both the detector and controller, or calls may be delayed longer than desired. The delay timer is only active when its assigned phase is not green.

Delay is commonly applied to (1) minor street stop bar detectors in exclusive right-turn lanes and (2) left-turn lane detectors. Delay is used to prevent erroneous calls from being registered in the controller (e.g., detectors in left-turn lanes may accidentally be traversed by vehicles on another phase or cut across by perpendicular left-turning vehicles) or to allow permitted movements to take place without calling a protected phase. Exhibit 6-23 illustrates delay assigned to stop bar detectors in an exclusive right-turn lane.

6.2.1.1 Typical Values for Delay

Appropriate values for delay depend on how long a vehicle is expected to occupy a detector before the vehicle leaves that detector. Exhibit 6-24 provides a summary of typical values for the following common applications:

- Delay may be used with stop bar, presence mode detection for turn movements from exclusive lanes. For right-turn-lane detection, delay should be considered when the capacity for right-turn-on-red exceeds the right-turn volume or a conflicting movement is on recall. If right-turn-on-red capacity is limited, then delay may only serve to degrade intersection efficiency by further delaying right-turning vehicles. The delay setting could range from 8 to 12 seconds, with the larger values used for higher cross-street volumes (*10*).
- Delay may also be used to prevent an erroneous call from being registered in the controller if vehicles tend to traverse over another phase's detection zone. For example, left-turning vehicles often cut across the perpendicular left-turn lane at the end of their turning movement. A detector delay coupled with non-locking memory would prevent a call from being placed for the unoccupied detector. The delay value depends on how long a detector is erroneously occupied. Typically, values range from 2 to 5 seconds.

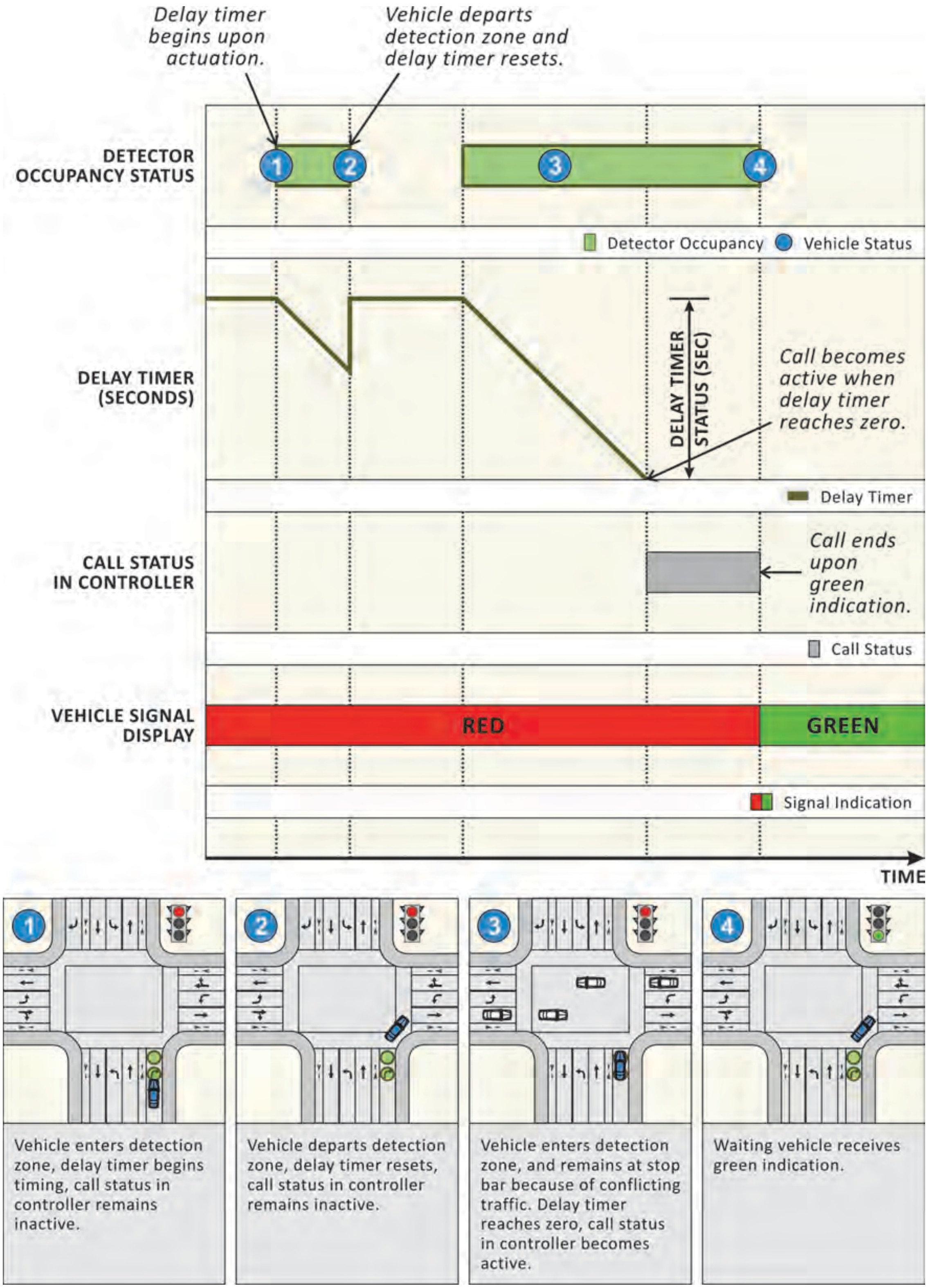


Exhibit 6-23 Delay Timer

Exhibit 6-24 Typical Values for Delay

Common Applications for Delay	Applicable Detectors	Delay (Seconds)
Right-Turn-on-Red in Right-Turn Lanes	Minor street stop bar detectors in exclusive right-turn lanes	8 to 12
Erroneous Call Prevention for Left-Turn Lanes	Stop bar detectors in left-turn lanes	2 to 5

6.2.2 Extend Time (Carry-Over or Stretch Time)

Extend time (also called carry-over or stretch time) is a method used to increase the duration of an actuation on a detector. Much like delay, extend time can be programmed on the detector or in the controller. Signal timing practitioners should make sure that extend time is not programmed both on the detector and in the controller, or calls may be extended longer than desired. The extend timer begins the instant the actuation channel input is inactive. Thus, an actuation that is 1 second in duration on the detector input can be extended to 3 seconds if the extend parameter is set to 2 seconds. This process is illustrated in Exhibit 6-25.

The extend parameter is typically used with detection designs that combine multiple setback detectors on high-speed approaches. The objective when used on high-speed approaches (with multiple detectors per lane) is to extend the green interval, allowing just enough time for a vehicle approaching the intersection to reach the next downstream detector. The vehicle will then be able to place a new call for green extension if traveling at the assumed speed.

This feature can also be used to provide lane-by-lane passage time. Instead of the passage time being controlled by the phase table, the detector times the extend parameter, allowing each lane to be timed separately. If stop bar detection is used, it should use the detector disconnect feature upon gap out. This allows the standing queue to discharge, but prevents the phase from being continually extended.

6.2.2.1 Typical Values for Extend Time with Multiple Detectors per Lane

The extend time assigned to a detector is dependent on the approach speed, detector size, and the distance between the subject detector and the next downstream detector. In order to maintain efficiency, typical values range up to 2 seconds. More information about specific extend time values can be found in the *Manual of Traffic Detector Design* (*11*).

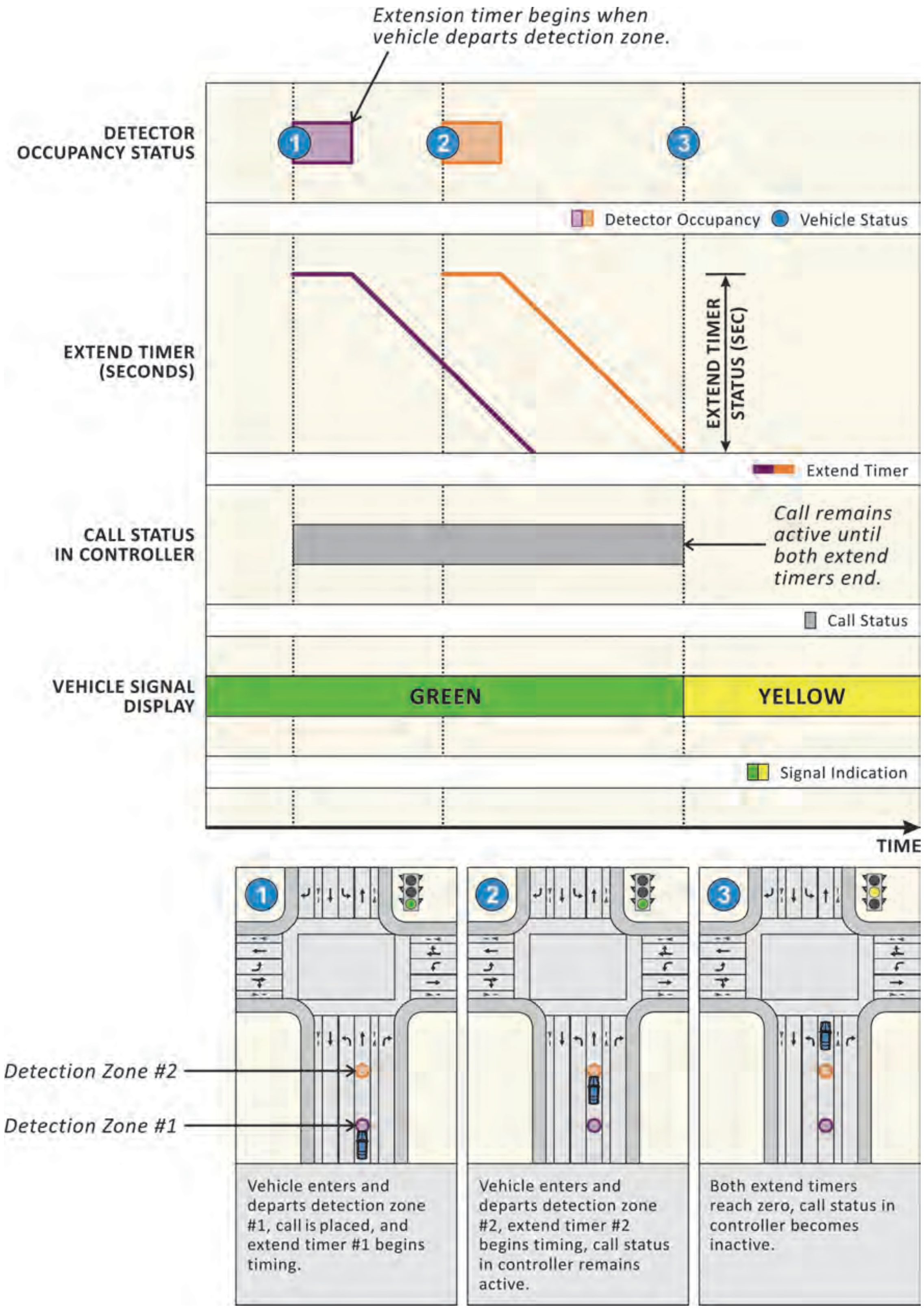


Exhibit 6-25 Extend Timer

6.2.3 Detector Switching

Detector switching is a common detector function in traffic signal controllers that allows detectors to call and extend one phase (extend phase) and then send calls to another phase (switch phase) once the extend phase ends. Detector switching allows the programmed switch phase to be extended after the extend phase terminates. It is typically only effective when the switch phase is green. Also, detector switching does not typically switch phase calls.

Detector switching is commonly used on left-turn lane detectors under protected-permitted operations. Vehicles detected in left-turn lanes are switched to extend the through phase during the permitted portion of the phase, in order to provide more time for vehicles making left-turn movements. Exhibit 6-26 provides a summary of typical detector switching settings at a standard eight-phase intersection with protected-permitted operations on all approaches.

Exhibit 6-26 Typical Settings for Detector Switching under Protected-Permitted Operations

Phase	Extend Phase	Switch Phase (with Five-Section Signal Head Protected-Permitted Operations)	Switch Phase (with FYA Protected-Permitted Operations)
1	1	6	2
3	3	8	4
5	5	2	6
7	7	4	8

6.3 TIME-OF-DAY PLANS

Most controllers allow signal timing parameter values to vary by time of day, week, or year. The schedule for these timing plan changes is referred to as the time-of-day plan. It provides information to the controller about which set of signal timing parameters should be used depending on the established schedule. Time-of-day plans should be developed for specific outcomes and can be used to help an agency meet its objectives during different time periods.

Determining when traffic conditions typically change can help a practitioner decide how many time-of-day plans are required. Most intersections experience a peak period during the morning and evening, so the morning and evening peaks may have different timing plans (with different signal timing values) than the rest of the day. For example, the maximum green on Phase 2 during a weekday may be programmed for 20 seconds most of the day and 30 seconds during the peak periods.

In addition to timing plan changes throughout a day, separate plans may be warranted for weekdays, weekends, or even specific days (i.e., Friday on a holiday weekend) based on variation in travel patterns and volumes. For example, tourist areas that experience a large fluctuation in traffic flow throughout the year may find it beneficial to develop timing plans for the tourist and non-tourist seasons. The practitioner will need to verify which parameters can be changed by time of day in the controller, but should consider different values when selecting the signal timing parameters explained throughout this chapter.

6.4 REFERENCES

1. *Uniform Vehicle Code*, Millennium Edition. National Committee on Uniform Traffic Laws and Ordinances, Alexandria, Virginia, 2000.
2. *Manual on Uniform Traffic Control Devices for Streets and Highways,* 2009 Edition. United States Department of Transportation, Federal Highway Administration, Washington, D.C., 2009.
3. *Determining Vehicle Signal Change and Clearance Intervals.* Institute of Transportation Engineers, Washington, D.C., 1994.
4. *Traffic Engineering Handbook*, 6th Edition. Institute of Transportation Engineers, Washington, D.C., 2009.
5. *Traffic Control Devices Handbook*. Institute of Transportation Engineers, Washington, D.C., 2001.
6. McGee, Sr., H., K. Moriarty, K. Eccles, M. Liu, T. Gates, and R. Retting. *NCHRP Report 731: Guidelines for Timing Yellow and All-Red Intervals at Signalized Intersections*. Transportation Research Board of the National Academies, Washington, D.C., 2012.
7. Tarnoff, P. J. Traffic Signal Clearance Intervals. In *ITE Journal*, Institute of Transportation Engineers, Washington, D.C., 2004, pp. 20–24.
8. Souleyrette, R. R., M. M. O'Brien, T. McDonald, H. Preston, and R. Storm. *Effectiveness of All-Red Clearance Interval on Intersection Crashes*. Report MN/RC-2004-26, Minnesota Department of Transportation, 2004.
9. *California Manual on Uniform Traffic Control Devices,* 2012 Edition. California Department of Transportation, Sacramento, California, 2012.
10. Bonneson, J. A., and M. Abbas. *Intersection Video Detection Manual*. Report FHWA/TX-03/4285-2, Texas Department of Transportation, Austin, Texas, 2002.
11. Bonneson, J. A., and P. T. McCoy. *Manual of Traffic Detector Design*, 2nd Edition. Institute of Transportation Engineers, Washington D.C., 2005.

CHAPTER 7
SYSTEM/COORDINATED TIMING

CONTENTS

LIST OF EXHIBITS

CHAPTER 7. SYSTEM/COORDINATED TIMING

Progression is the movement of users along a designated route in a manner that minimizes stops.

Instead of each signal operating independently, coordination allows signals to operate as a group, thereby synchronizing movements and allowing for better progression. Chapter 7 explains how the basic timing parameters introduced in Chapter 6 can be used in conjunction with coordinated features to operate a group of signals. In addition to providing guidance on the signal timing parameters (and typical values) that must be defined for coordination, Chapter 7 also highlights the advantages, disadvantages, and complexities associated with such a system.

Traditional coordination requires that cycle length, splits, and offsets be defined.

Traditional coordination adds a layer of signal controller logic to the basic actuated logic described in Chapter 6. This system control method requires the practitioner to define a consistent cycle length for the corridor, as well as splits and offsets for each timing plan. Splits control the amount of time given to each phase in a cycle, and offsets control the time relationship between intersections. Both are described in detail in the following sections,

Traditional coordination will continue to be the dominant form of control for the foreseeable future, but some advanced systems (and likely future systems) will deviate somewhat from the traditional cycle/split/offset model. Some adaptive systems are able to maintain coordination without having to explicitly define cycle length, splits, and offsets; they are able to adjust timing values based on measured conditions. More information about advanced systems and features is available in Chapter 9.

7.1 APPLICATION OF A COORDINATED SYSTEM

The decision to use coordination should be influenced by a variety of factors, but most importantly, a practitioner should consider the operating environment, users, and appropriate user priorities. Coordination should be used to meet specific objectives; it is not appropriate in every situation. While coordination can reduce travel time, stops, delay, and queues for the coordinated movements, there may be consequences for the uncoordinated movements.

When intersections are close together, it may be advantageous to coordinate them to accommodate arriving platoons.

Information presented in the Federal Highway Administration (FHWA) report *Signal Timing on a Shoestring* (*1*) reveals that both simple and complex procedures can be used to identify which intersections to coordinate. In general, when intersections are close together (i.e., within one-half mile of each other), it is advantageous to coordinate them, particularly if volumes between the intersections are large. At greater distances (i.e., one-half mile or greater), the traffic volumes and potential for platoons should be further reviewed to determine if coordination would benefit system operations. If there is minimal traffic variation between intersections (e.g., caused by vehicles turning into or out of accesses), coordination at distances of a mile or more may be appropriate.

If corridor progression is the defined objective, the need for coordination can often be identified through observation of the traffic flow arriving from upstream intersections. If arriving traffic includes platoons that have been formed by the release of vehicles from an upstream intersection, coordination may provide desired progression benefits. On the other hand, if vehicle arrivals (a) tend to be random and are unrelated to the upstream intersection operations or (b) are broken up by accelerating/decelerating vehicles due to driveways or bus blockages, coordination may provide little benefit to the system.

The costs and benefits of using coordination should also be weighed carefully against a community's multimodal goals. Many of the complaints from citizens related to the use of coordination are about the added delay and lack of responsiveness to their demand, particularly when there is no demand on conflicting movements. Delay for minor movements will be further exacerbated if the selected cycle length is unnecessarily long or the coordination plan is operating when traffic volumes are lower than is typical (e.g., during holidays that fall on a weekday).

7.2 COORDINATION PLANNING USING A TIME-SPACE DIAGRAM

As mentioned previously, traditional coordination uses the cycle/split/offset method to progress traffic. Various software and manual tools can be used to optimize this combination of signal timing parameters, but the practitioner should always ensure that the "optimized" results reflect the outcome based process (described in Chapter 3). (Typical software programs have a focus on ***system*** vehicle delay, which may ***not*** be an appropriate operational objective.) One tool that is often used to review coordination is the time-space diagram, which can be drawn manually or created using software. Basic elements of the time-space diagram are explained in this section, as well as how it can be used to evaluate coordinated timing plans.

Time-space diagrams are a visual tool that can be used to assess coordination strategies and evaluate timing plans before field implementation.

Time-space diagrams are a visual tool that practitioners commonly use to analyze coordination strategies and modify timing plans. The diagrams illustrate the relationship among intersection spacing, signal timing, and vehicle movement. When used with some software programs, they can also be used to derive vehicle-based performance measures (also known as measures of effectiveness or MOEs), such as stops, vehicles arriving on green, and queue lengths. Coordinated parameter definitions and guidance are provided throughout the following sections, and many of them are explained using components of the time-space diagram in Exhibit 7-1.

7.2.1 Diagram Axes

Time-space diagrams are drawn with time on the horizontal axis and distance (from a reference point) on the vertical axis. In Exhibit 7-1, distance has been depicted using an aerial view of three signalized intersections. For ease of initially displaying relationships, the minor streets are all one-way (which is not typical). It is very important that time-space diagrams are drawn to scale so that an accurate relationship between time and distance is maintained for vehicle progression calculations.

The time-space diagram examples throughout this chapter assume that the corridor being analyzed is a north-south roadway. However, time-space diagrams can be used to analyze corridors traveling in any direction. The practitioner must simply assign a reference point for the distance axis. For example, an east-west corridor could be shown with the western-most intersection at the top of the vertical axis and the eastern-most intersection at the bottom.

7.2.2 Master and Local Clocks

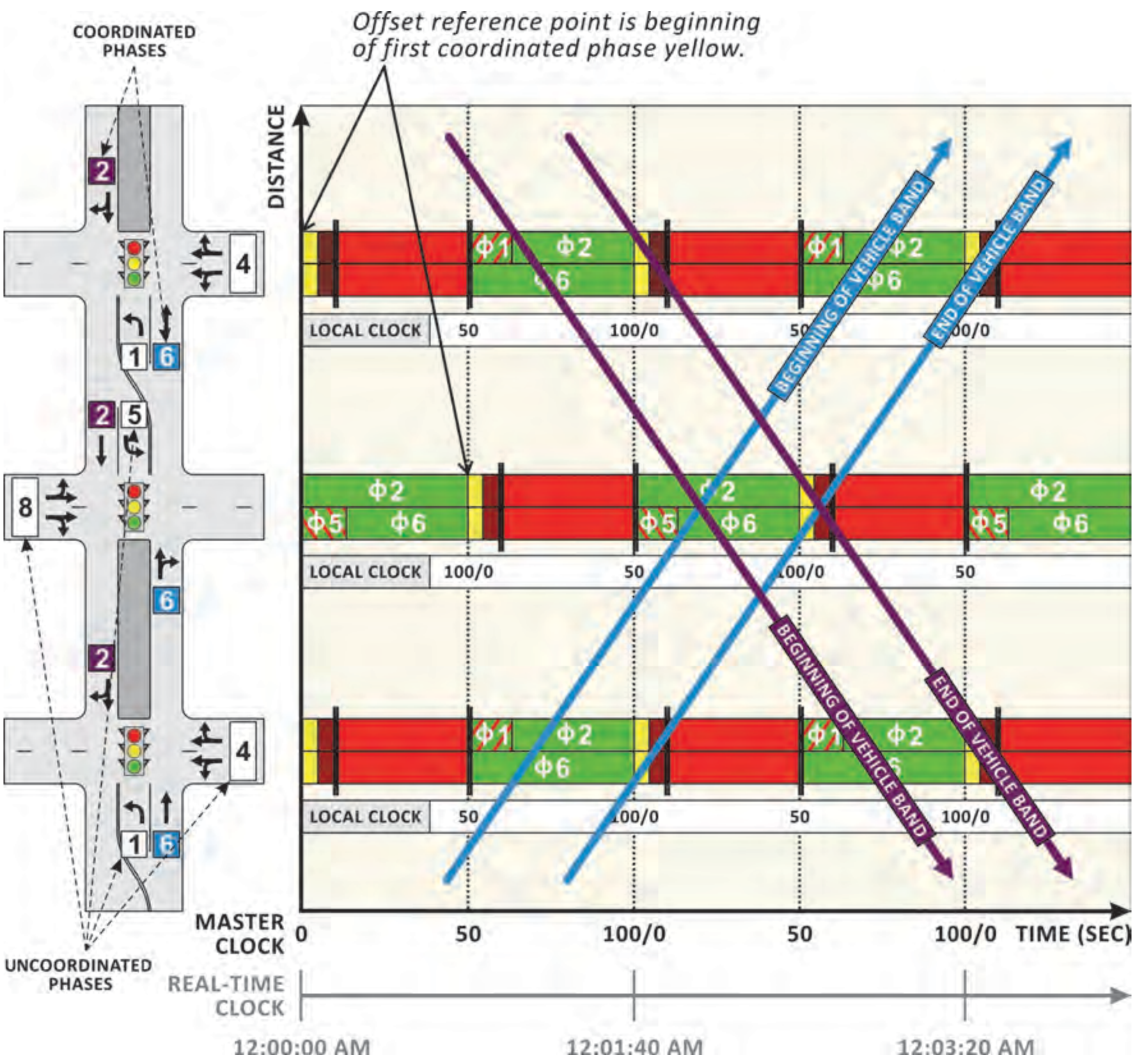
Two types of clocks are running in the background during coordinated operations—the system master clock (i.e., field master, master controller, or central system) and the local (intersection controller) clock. The master clock is the background timing mechanism within the controller logic to which each local controller is referenced in

The relationship between the master clock and the local (intersection controller) clocks is used to establish coordinated operations.

order to establish an offset between intersections (see Section 7.3.9). Both clocks must have the same time in order for coordination to work.

In the time-space diagrams in this chapter, the master clock is shown at the bottom of the diagram on the horizontal axis. The local clocks (if shown) are depicted below the ring-and-barrier diagrams associated with each intersection. In all exhibits in this manual, the offset reference point (see Section 7.4.8) is set to the beginning of first coordinated phase yellow (end of green), which is the beginning (zero point) of the local clock cycle. To help illustrate the relationship between the clocks, the master clock and local clocks in Exhibit 7-1 are shown in relation to a real-time clock below the horizontal axis. While the real-time clock shows times of day (note Exhibit 7-1 uses a sync reference of midnight, see Section 7.3.7), the master clock and local clocks show time relative to cycle length (in seconds).

Exhibit 7-1 Basic Time-Space Diagram



7.2.3 Signal Timing Graphics

In the main portion of the diagram, signal timing information is provided for each intersection. Ring-and-barrier diagrams (explained in Chapter 5) depict the signal indications for the coordinated movements. In Exhibit 7-1, Phases 2 and 6 are the coordinated phases (which is typical but not an absolute requirement), and the ring-and-barrier diagrams show when those phases are receiving a green, yellow, or red indication. The dark red indicates the red clearance interval, and the bright red indicates when other uncoordinated phases (in this case, Phases 4 and 8) are receiving their green, yellow, and red clearance indications. Actual green times ***may*** differ from those shown in the ring-and-barrier diagrams because of gap outs and/or uncoordinated phases being skipped. Note that ring-and-barrier diagrams may show both rings (as in Exhibit 7-1) or a combined graphic that depicts both rings together (explained in detail in the following section).

Ring-and-barrier diagrams graphically represent phases and their relationships at an intersection.

7.2.4 Signal Timing Graphics for Left-Turn Phases

Signal timing information for left-turn phases (adjacent to the coordinated movements) is also depicted in a time-space diagram (see Exhibit 7-2). Protected left-turn movements are represented by hatching that is in the same direction as the through movement in the other ring. For example, with Phase 1 accommodating a northbound left-turn movement at the northern-most intersection, the diagonal hatching for Phase 1 is in the same direction as a northbound Phase 6 through vehicle trajectory. Because a southbound Phase 2 through vehicle trajectory would be in the opposite direction, a practitioner can easily identify that a southbound through vehicle would not be able to proceed through the intersection when Phase 1 is being served.

A vehicle trajectory is a representation of a vehicle traveling along the corridor that depicts both distance and time; it is explained in detail in the next section.

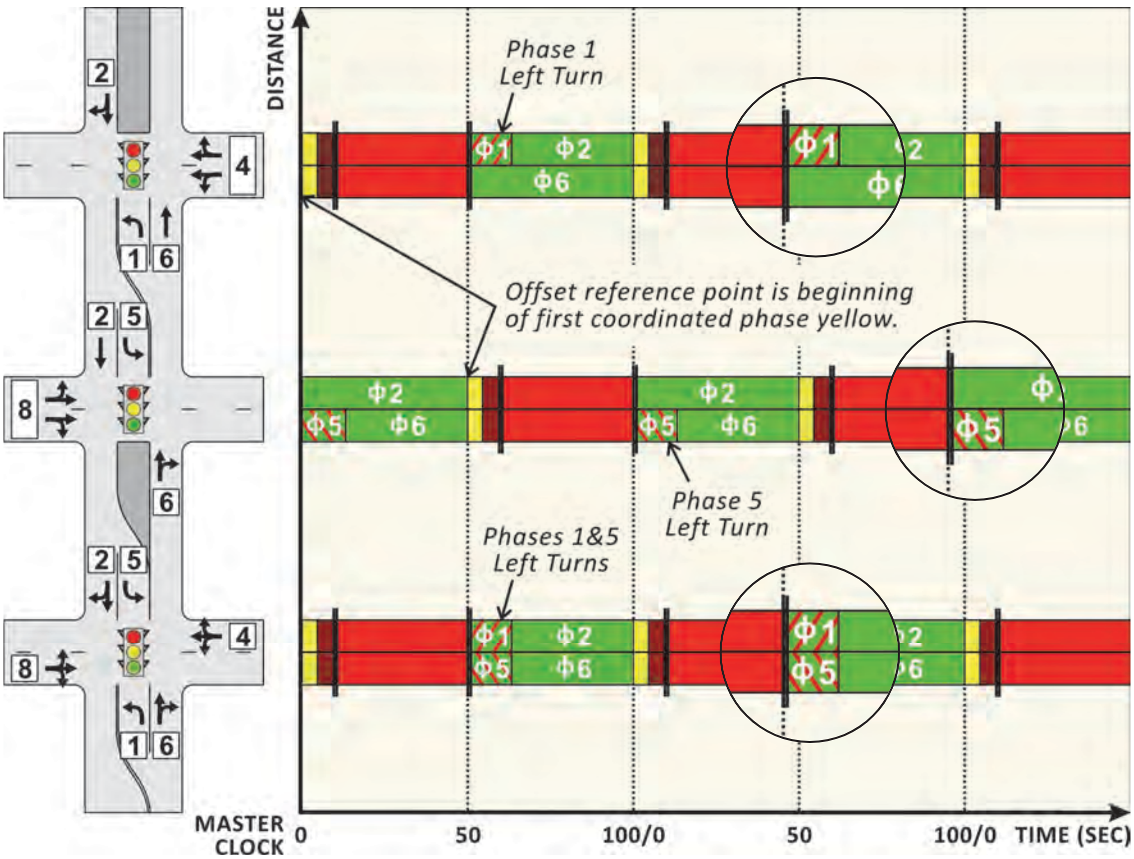
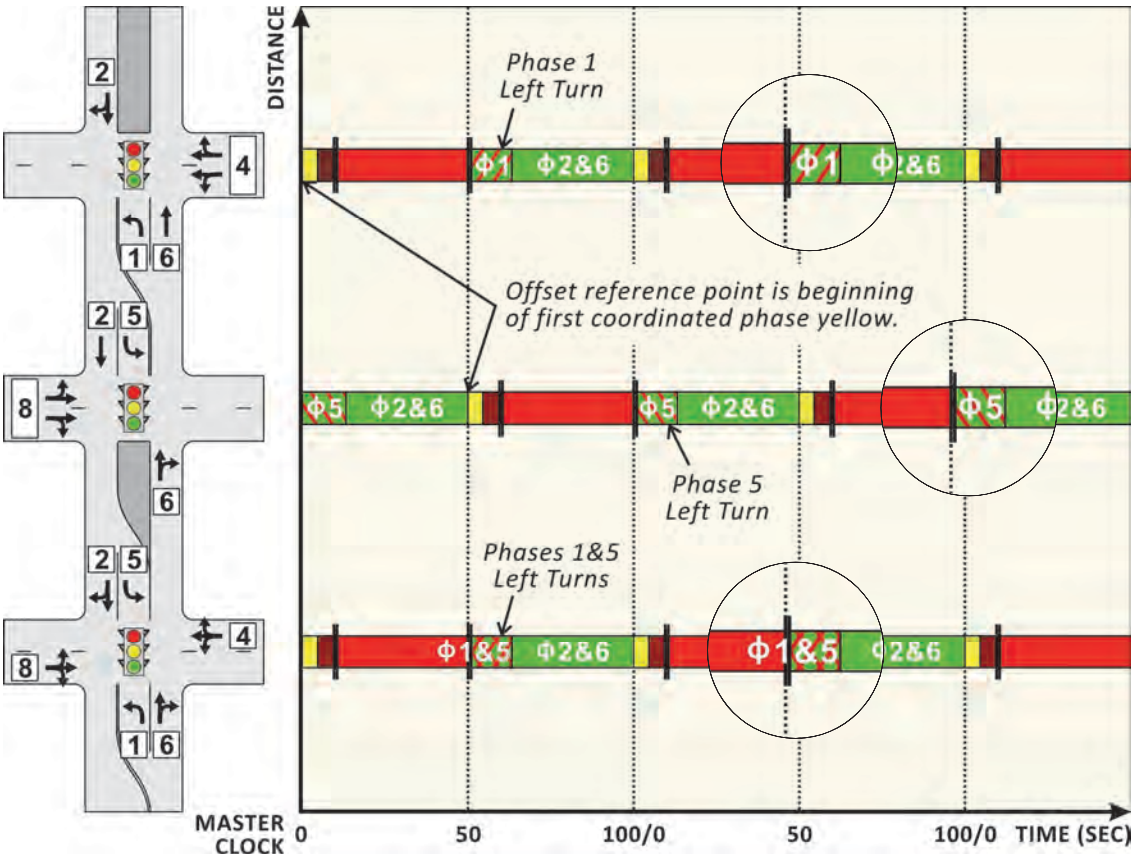


Exhibit 7-2 Time-Space Diagram Left-Turn Phasing (Separate Rings)

When two protected left-turn phases occur at the same time (as shown at the southern-most intersection of Exhibit 7-2), the hatching shows a period of time when through movements are not possible in either direction. The two rings may not always be shown separately; in a combined ring-and-barrier diagram (see Exhibit 7-3), left-turn phases that occur at the same time are shown using a crisscross pattern (as highlighted at the southern-most intersection in Exhibit 7-3). These hatching conventions for left-turn phasing should make it easier for a practitioner to know at what point in the cycle a vehicle can progress.

Exhibit 7-3 Time-Space Diagram Left-Turn Phasing (Combined Rings)



7.2.5 Vehicle Trajectories

One of the most important components of a time-space diagram is the representation of vehicles traveling along the corridor. Vehicles are depicted using trajectory lines shown on top of the ring-and-barrier portion of the diagram. In Exhibit 7-1, they represent vehicles moving either north or south (toward the top or bottom of the diagram, respectively) as time increases (from left to right). (Note that if the distance axis represented an east/west street, the vehicle trajectories would depict eastbound/westbound vehicles.) Assuming that Exhibit 7-1 depicts a corridor with the northern-most intersection shown at the top of the diagram, the diagonal blue lines going from bottom/left to top/right represent northbound vehicles, and the diagonal purple lines going from top/left to bottom/right represent southbound vehicles. (Note that trajectory lines are not typically color-coded; blue and purple are used for example clarity only.)

The time-space diagram examples throughout this chapter show vehicle trajectories crossing the ring-and-barrier diagrams at the stop bar locations associated with the direction of travel (as shown in Exhibit 7-4). In other words, the northbound vehicles are drawn entering the intersections according to the signal indications at the bottom of the ring-and-barrier diagrams, and the southbound vehicles are drawn entering the intersections according to the signal indications at the top of the ring-and-barrier diagrams. With a combined ring structure (as shown in Exhibit 7-3), a thinner ring-and-barrier diagram can be used. Thinner ring-and-barrier diagrams make the "point of entry" conventions shown in Exhibit 7-4 of less significance. However, the remainder of this chapter uses ring-and-barrier diagrams with separate rings, in order to clearly demonstrate coordinated system concepts for both coordinated phases.

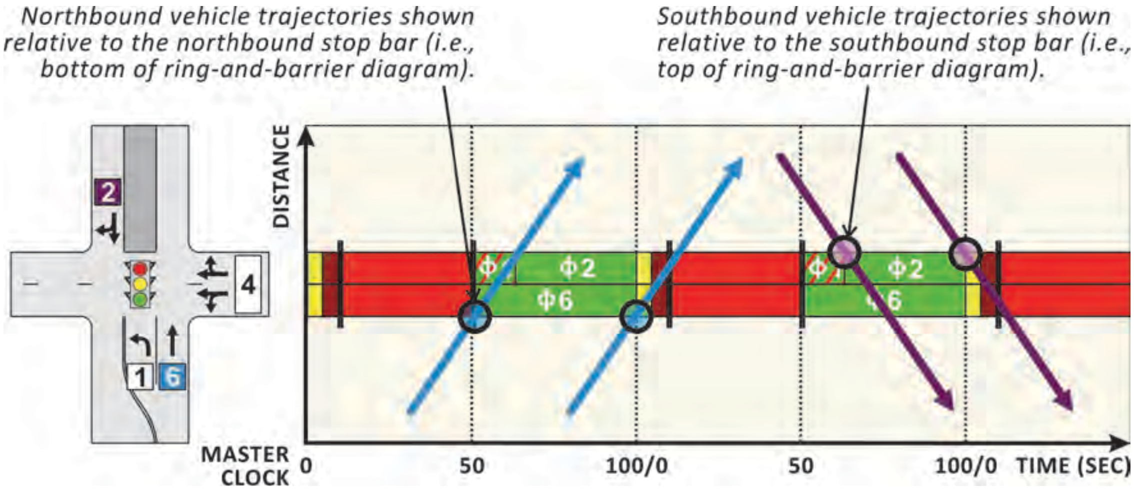


Exhibit 7-4 Vehicle Trajectory References

Because of the relationship between distance and time established in a time-space diagram, vehicle trajectories can be used to illustrate the speed that vehicles can progress along a corridor. The distance traveled by a vehicle divided by the elapsed time (distance divided by time, or change in y divided by change in x) represents the speed (see Exhibit 7-5). The assumed speed for coordination (i.e., the progression speed) may be the speed limit, the 85th-percentile speed, or a desired speed set by the practitioner.

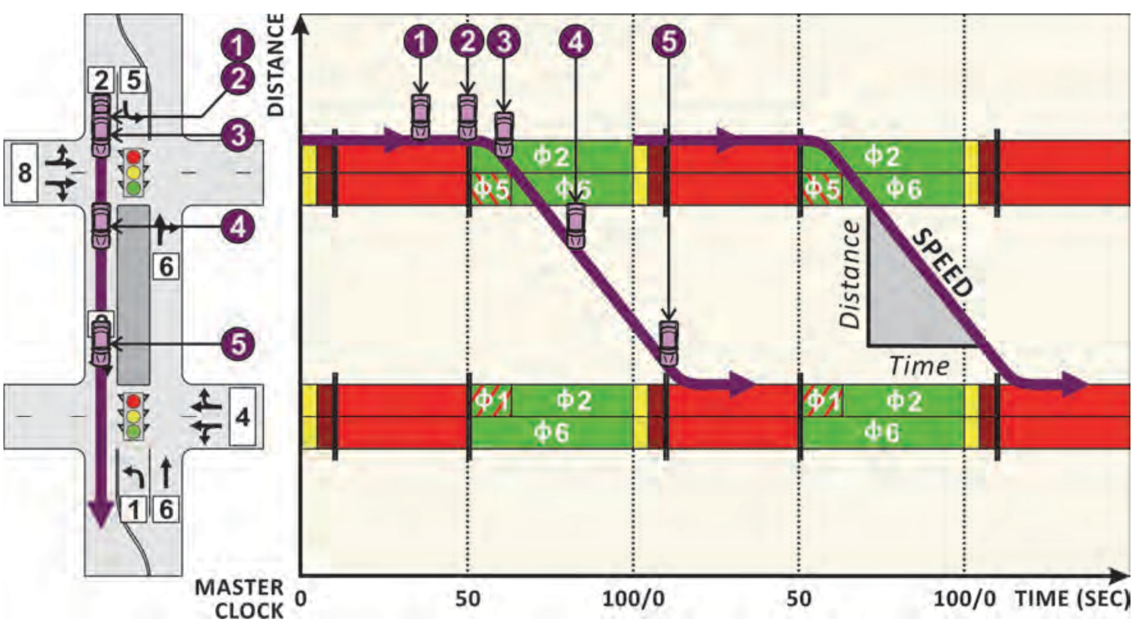


Exhibit 7-5 Vehicle Conditions on a Time-Space Diagram Trajectory

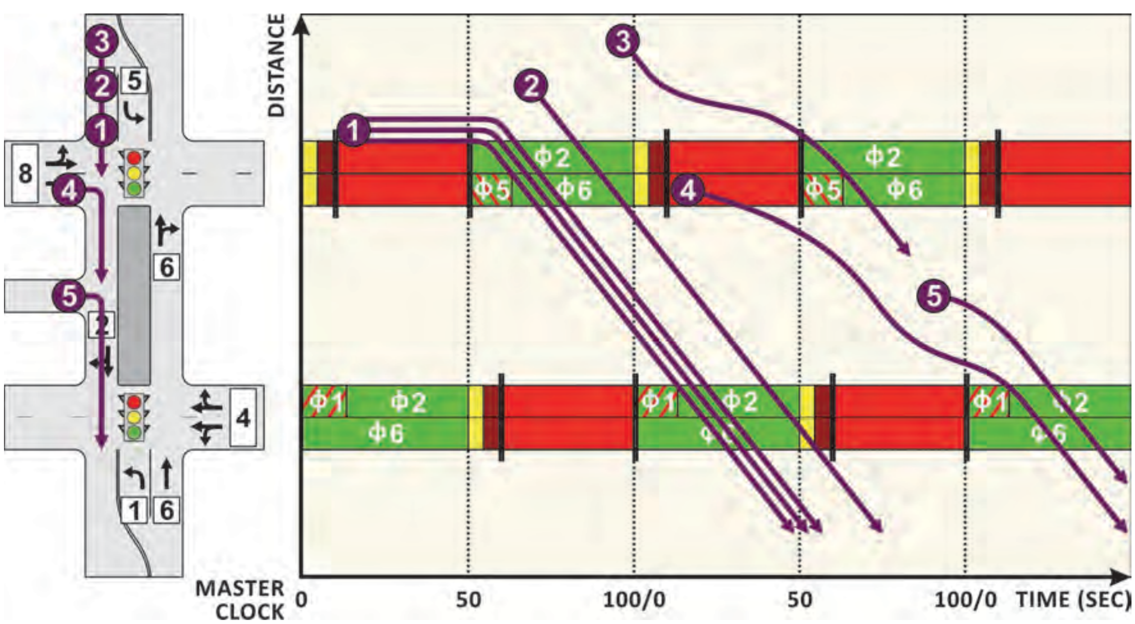
Exhibit 7-5 also illustrates the five different conditions a driver can experience when traveling between two signalized intersections. Each condition is described in detail, as it relates to the vehicle trajectory, in Exhibit 7-6. In this example, the vehicle trajectory shows that progression could be improved for the southbound movement by adjusting the offset (i.e., time relationship between intersections) so that the vehicle would not have to stop at both intersections.

Exhibit 7-6 Vehicle Conditions Associated with a Trajectory

No.	Condition	Description	Time-Space Diagram Representation
1	Stopped	Vehicle can be initially delayed as a result of a red indication at the signal.	Horizontal line; no change in distance as time moves forward.
2	Perception/ Reaction Time	Drivers at the front of the queue must react to the green indication.	Horizontal line transitions to diagonal line; driver begins moving forward, increasing distance from the first intersection with time.
3	Acceleration	Vehicle accelerates up to the progression speed.	Diagonal line; vehicle moving forward with time.
4	Running Speed	Vehicle travels at the progression speed to the next intersection (often assumed to be the speed limit or an estimated progression speed).	Diagonal line; vehicle moving forward with time.
5	Stopping	Vehicle slows to stop due to red light (or stopped traffic).	Diagonal line transitions to a horizontal line.

Exhibit 7-7 shows a wider range of possible vehicle trajectories between two signalized intersections. Progression has improved over that in Exhibit 7-5, as vehicles can progress through both intersections (traveling at the progression speed) without stopping.

Exhibit 7-7 Example Vehicle Trajectories



The vehicle trajectories in Exhibit 7-7 (labeled 1 through 5) are described below:

1. Three vehicles travel from a stopped position at the first intersection through the downstream intersection.

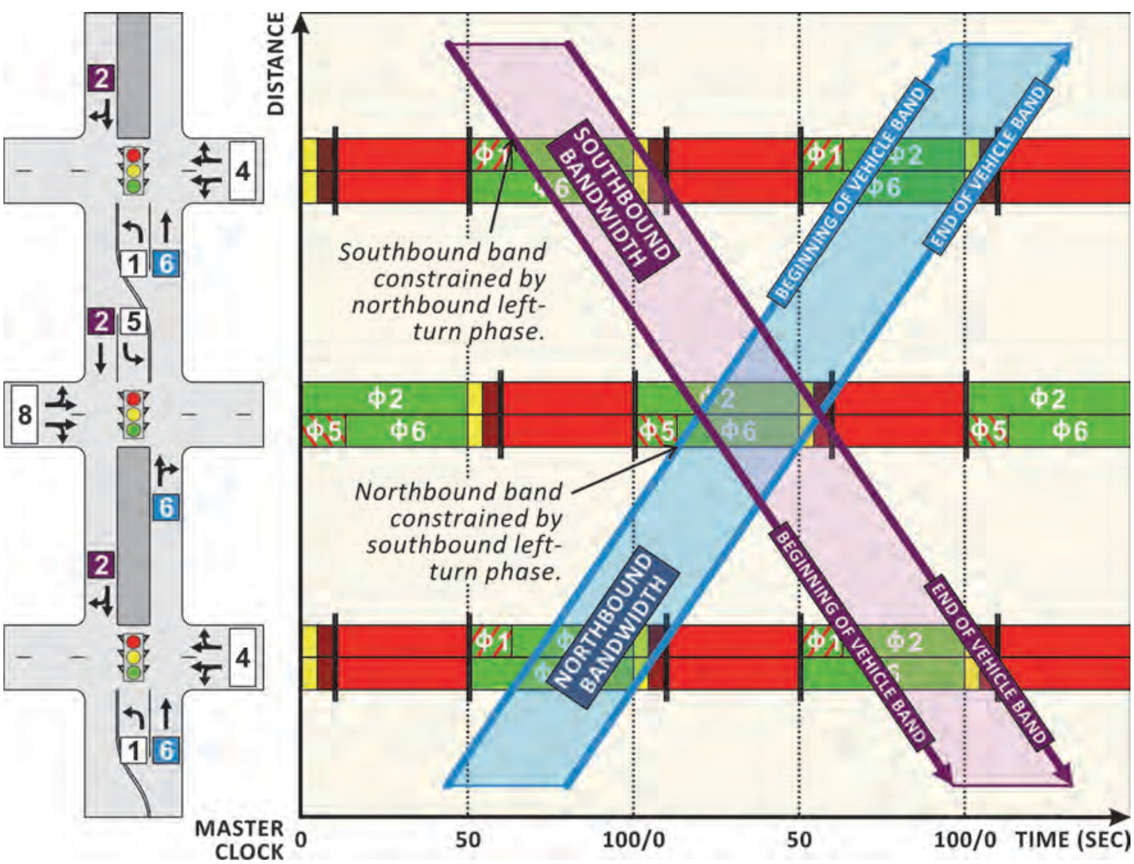
2. A vehicle traveling at the progression speed travels through both intersections without stopping.
3. A vehicle delayed at the upstream intersection (not coming to a complete stop) travels through the intersection upon the green indication.
4. A vehicle from the minor street travels to the downstream intersection. The vehicle enters the corridor when Phase 8 is receiving a green indication, which is a point in time when the coordinated phases are receiving a red indication. The vehicle has to slow as it approaches the downstream intersection, as shown by the more gradual slope of the trajectory.
5. A vehicle from a mid-block driveway travels through the downstream intersection.

7.2.6 Bandwidth

Section 7.2.5 stated that vehicle trajectories show the movement of vehicles as they progress along a coordinated corridor. However, there are select trajectories that illustrate vehicles traveling along a coordinated corridor without slowing or stopping. The first and last of these select vehicle trajectories outline the bandwidth (as shown in Exhibit 7-8). As the bandwidth gets wider, potential progression opportunities increase for vehicles traveling along the coordinated corridor. However, practitioners should always consider the system objectives. Wide bandwidths will only be effective if there are enough vehicles to utilize them.

Bandwidth represents progression opportunities and is illustrated on time-space diagrams as the shaded area between vehicle trajectories.

Exhibit 7-8 Bandwidth



Bandwidth is an ***ideal*** representation of progression, in that it does not explicitly account for vehicle acceleration from a stop, dispersion of vehicles as they travel from one intersection to the next, or queued vehicles at the downstream intersections. Actual bandwidth with actuated phases may vary due to phases gapping out (and not using their total split time). In addition, detector failure will limit the ability of bands to expand during light traffic. However, bandwidth is a useful tool for evaluating the relationship between coordinated parameters (including phase sequence, cycle length, splits, and offsets). A few important points to understand related to bandwidth are

- Bandwidth is dependent on the selected progression speed.
- Bandwidth can be different for each direction of travel.
- As additional intersections are added to a system, it is increasingly difficult to provide progression. It is sometimes better to break a long corridor into smaller segments (at convenient locations), particularly corridors with longer spacing between intersections. The segments can then be combined by selecting an offset relationship between the separately optimized segments, generally providing good progression in one direction. This approach is sometimes called a programmed stop. The stop location should be where there is adequate distance to provide storage without impacting the upstream intersection.
- During periods of oversaturated conditions (i.e., volume exceeds capacity), calculated bandwidth may never occur due to the influences of queuing. Strategies for mitigating oversaturated conditions are available in Chapter 12.
- Depending on the traffic volumes, timing plans that seek the greatest bandwidth may increase network delay and fuel consumption due to increased delay for uncoordinated movements.

Accommodating left-turning vehicles at signalized intersections is a balance among intersection safety, capacity, and delay.

Phase sequence can have a significant impact on bandwidth and should be evaluated when intersections are not equally spaced or have different offsets. In particular, leading or lagging the left-turn movements can greatly influence the progression opportunities for the through movements. For example, Exhibit 7-9 shows a corridor with leading left turns, which results in fairly small northbound and southbound bands.

Exhibit 7-10 shows how the bandwidth can be increased for the northbound movement by using lead-lag phasing at the middle intersection. The northbound through movement is able to utilize the additional green time at the beginning of the cycle. While not shown in Exhibit 7-10, the southbound bandwidth could also have been increased by changing the phase sequence at the other two intersections. The practitioner is cautioned that lagging lefts are less efficient (see Chapter 5) and should only be used where the increase in coordination performance is significant.

The opportunity for progression may be further enhanced by using protected-permitted lead-lag phasing, which is feasible with flashing yellow arrows (FYAs). (For more information on FYAs and potential conflicts, refer to Chapter 4.) This technique can reduce the time necessary for the protected interval, which restricts movement on the opposing through phase. Protected-permitted lead-lag left-turn phasing is especially effective for coordinated signals where the progressed platoons in each direction do not pass through the intersection at exactly the same time (as is the case for the examples in Exhibit 7-9 and Exhibit 7-10).

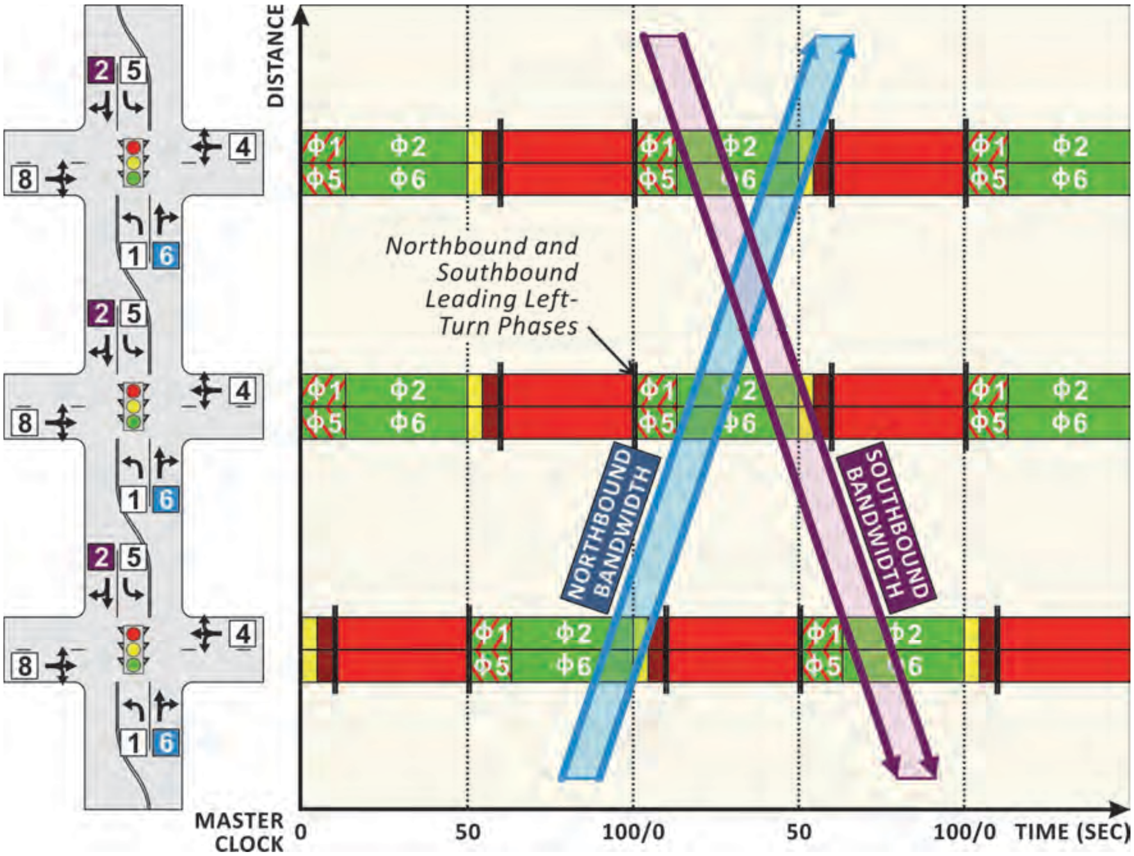


Exhibit 7-9 Phase Sequence Impacts on Bandwidth (Leading Left Turns)

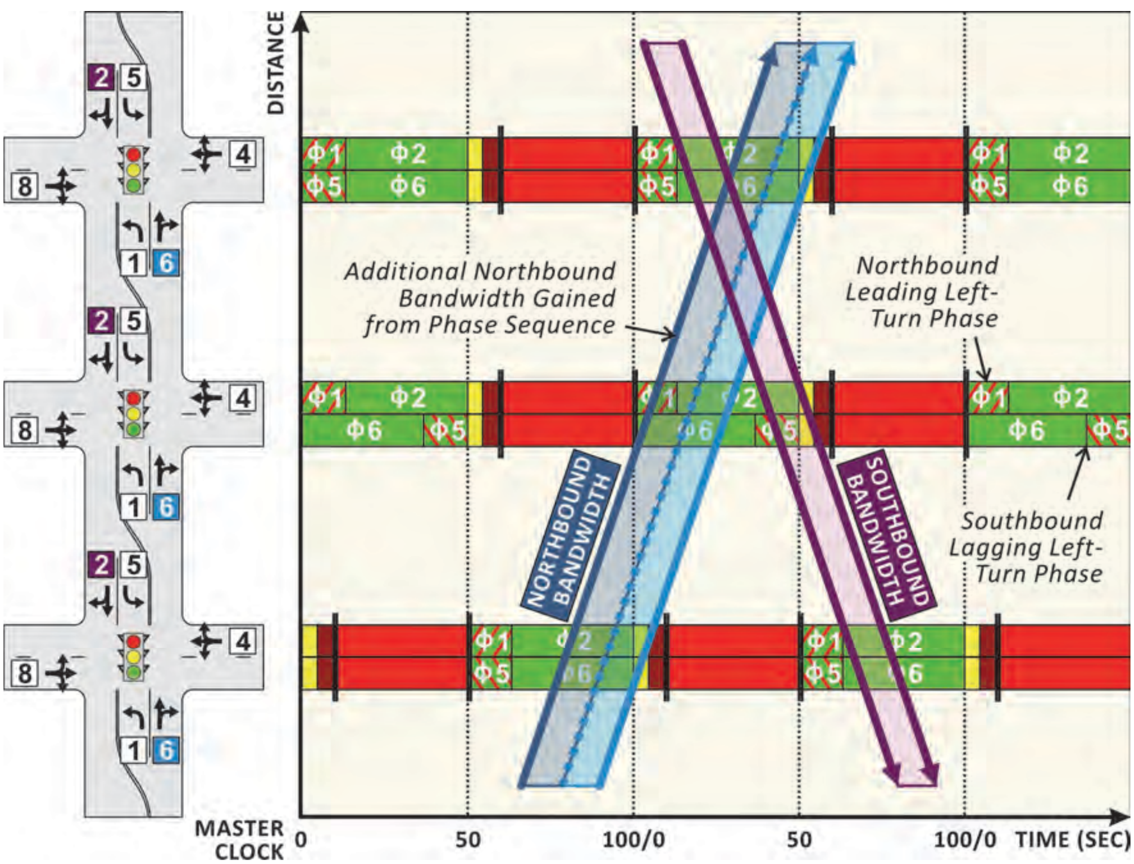


Exhibit 7-10 Phase Sequence Impacts on Bandwidth (Lead-Lag Left Turns at Middle Intersection)

7.3 INTRODUCTION TO COORDINATION PARAMETERS

As mentioned previously, there are several fundamental parameters that must be selected or programmed in order for a traditional coordinated signal system to operate. This section describes these parameters, often using pieces of the time-space diagram introduced in Exhibit 7-1. Guidance on parameter ***values*** is provided in Section 7.4.

While this manual will generally use National Transportation Communications for ITS Protocol (NTCIP) "NTCIP 1202—Object Definitions for ASC" (*2*) terminology and data standards (i.e., format of inputs), it will also highlight when alternative terms or formats are used and may be preferred. It is important to understand that the NTCIP does not define how a controller works. It defines the ***meaning*** of parameters that can be used in any NTCIP compliant controller. For example, the NTCIP term "pattern" refers to a unique set of coordination parameters—cycle, split, offset, and sequence.

7.3.1 Coordinated Phases

Coordination requires the designation of a phase or multiple phases as the coordinated phase(s). Coordinated phases are selected so that traffic flow can be prioritized between multiple intersections, typically to reduce stops and delay for the coordinated movements. As mentioned previously, time-space diagrams focus on the coordinated phases and show signal indications and vehicle movements relative to those phases. While the examples in this manual focus on vehicle coordination, strategies to coordinate pedestrian, bicycle, and transit movements can also be implemented based on local desired outcomes.

Coordinated phases are distinguished from uncoordinated phases because they always receive a minimum amount of assigned green time every cycle. While it is possible to have an actuated portion of the coordinated phase(s) (a process that is discussed further in Section 7.5.2), there is always a non-actuated interval that is "guaranteed" every cycle. With a consistent cycle length, the guaranteed green interval can be used to maintain the coordinated relationship between intersections.

7.3.2 Cycle Length

Cycle length is the time required for a complete sequence of signal phases at an intersection. For an actuated control system, a complete cycle is dependent on the presence of calls on all phases. For a pretimed control system, a complete cycle can be recognized by a complete sequence of signal indications. In the time-space diagram shown in Exhibit 7-11, a cycle is depicted as one set of green, yellow change, red clearance, and red indications for the coordinated phases. The master clock (depicted on the horizontal axis) is shown relative to the programmed cycle length, which is 100 seconds in this example. As discussed in Section 7.2.2, the master clock is referenced to a real-time clock (typically at midnight or another time in the early morning).

All of the intersections included in a coordination plan should have the same cycle length, in order to maintain a consistent time-based relationship between intersections. One exception is an intersection that "double cycles," serving phases twice as often as the other intersections in the system. In that case, the cycle length would be half the cycle length used at other intersections (as highlighted in Exhibit 7-11).

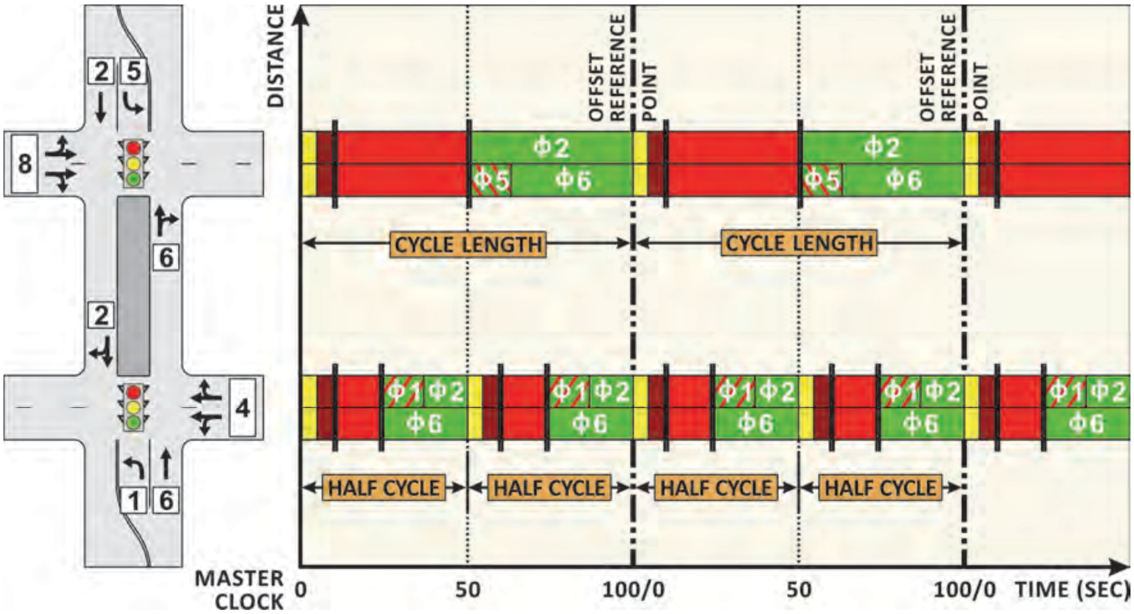


Exhibit 7-11 Cycle Length

7.3.3 Splits

Splits are the portion of the cycle allocated to each phase (green time plus clearances). Coordinated splits are selected based on the intersection phasing and expected demand. They can be expressed either in seconds (as outlined in NTCIP 1202, *2*) or percentages of the cycle length. Split times include the green time, yellow change interval, and red clearance interval associated with a phase. Exhibit 7-12 highlights the splits associated with the coordinated phases. For implementation in a signal controller, the sum of the phase splits must generally be equal to (or less than) the cycle length if measured in seconds (or 100 percent if measured as a percent).

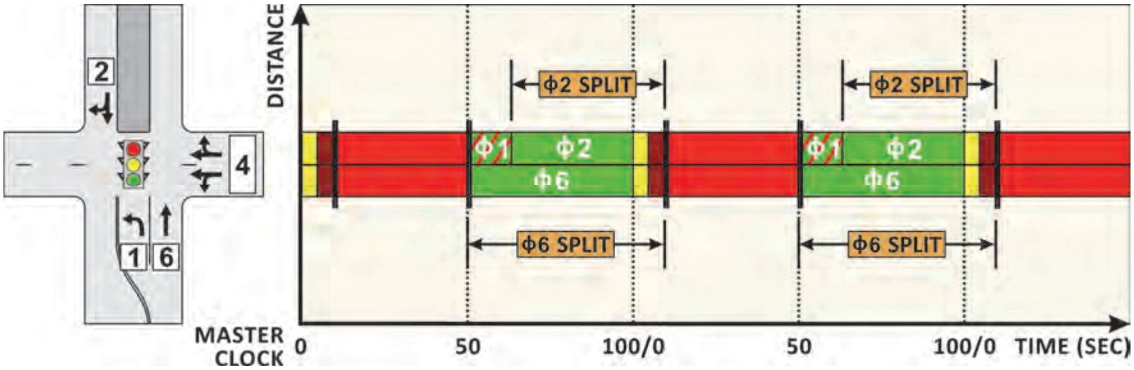


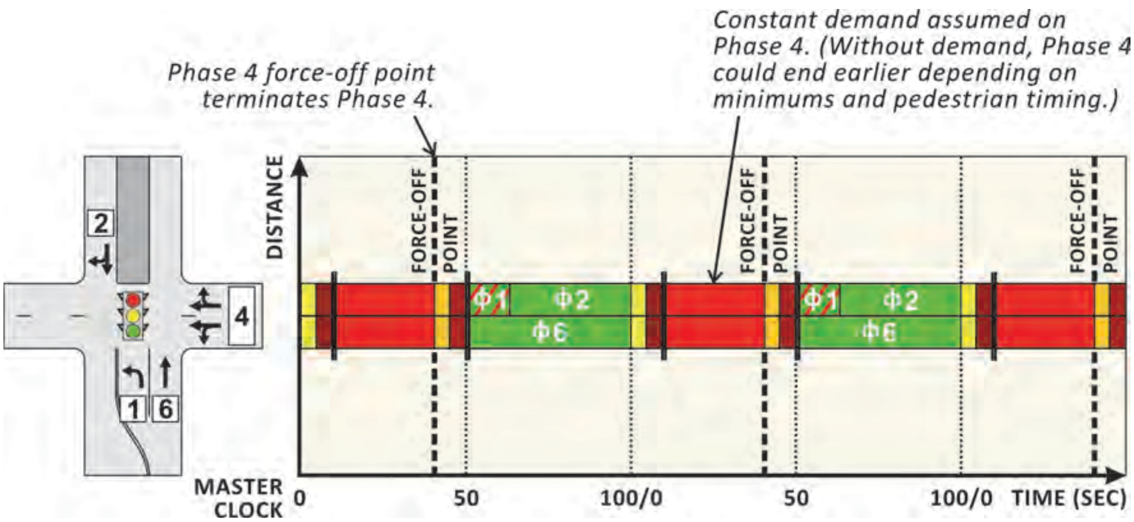
Exhibit 7-12 Splits

There are some controller firmware products that will allow splits to be shorter than pedestrian requirements. This is intended for situations in which there are few pedestrian calls. Full pedestrian times must always be served if called, so if splits do not include adequate time for pedestrian movements, the controller may need to transition in the next cycle to maintain coordination. More information on transition modes and pedestrian impacts is available in Section 7.5.

7.3.4 Force-Offs

Force-offs are used to enforce phase splits. They define the time during the cycle at which uncoordinated phases must end, even if there is continued demand, and ensure that control returns to the coordinated phase(s) no later than the programmed time, as illustrated in Exhibit 7-13. (Note that Phases 4 and 8 are shown in red because the time-space diagram highlights the signal status of the coordinated corridor.) Most modern controllers calculate force-off times based on the programmed split times (or split percentages). It should be noted that force-offs cannot override minimum times or clearance intervals. If force-offs are violated (i.e., the phase continues to serve minimum times or clearance intervals), the controller may have to re-sync to get back into coordination, unless following uncoordinated phases can give up time to allow the coordinated phase(s) to be served at their assigned time.

Exhibit 7-13 Force-Offs



There are two types of force-offs—floating and fixed—that determine how unused, actuated green time on the uncoordinated movements is shared with subsequent movements (as shown in Exhibit 7-14). ***Floating*** force-offs limit a phase to the assigned split time that was programmed into the controller. If a phase does not use all of its allocated time, then that movement cannot give the extra time to the next uncoordinated phase(s); all extra time is inherited by the coordinated phase(s). A ***fixed*** force-off maintains a phase's force-off point within the cycle. If a previous uncoordinated phase ends early, any following phase may use the extra time up to that phase's force-off point.

Exhibit 7-14 contrasts floating and fixed force-offs (in a single-ring example with Phase 2 being the coordinated phase). The top row illustrates a scenario in which demand equals or exceeds the allotted green time for all movements, and each phase is terminated at its respective programmed force-off point (sometimes referred to as "maxing out"). In other words, there is enough demand that none of the uncoordinated phases (Phases 3, 4, and 1) are able to end early. Note that instead of a force-off point, Phase 2 ends at a yield point. Yield points are explained in Section 7.3.6.

The next two rows illustrate the floating and fixed force-off concepts using a scenario in which there is unused green time on the uncoordinated movements. To better illustrate the difference between the two concepts, the demand affecting each phase (in seconds) is shown directly above each ring-and-barrier diagram. In this case, Phase 4 desires 30 seconds when its split time is 25 seconds. Under fixed force-offs, Phase 4 gets more time because Phase 3 only used 10 seconds of its allocated 25 seconds. Under floating force-offs, Phase 4 can never inherit any of Phase 3's unused time.

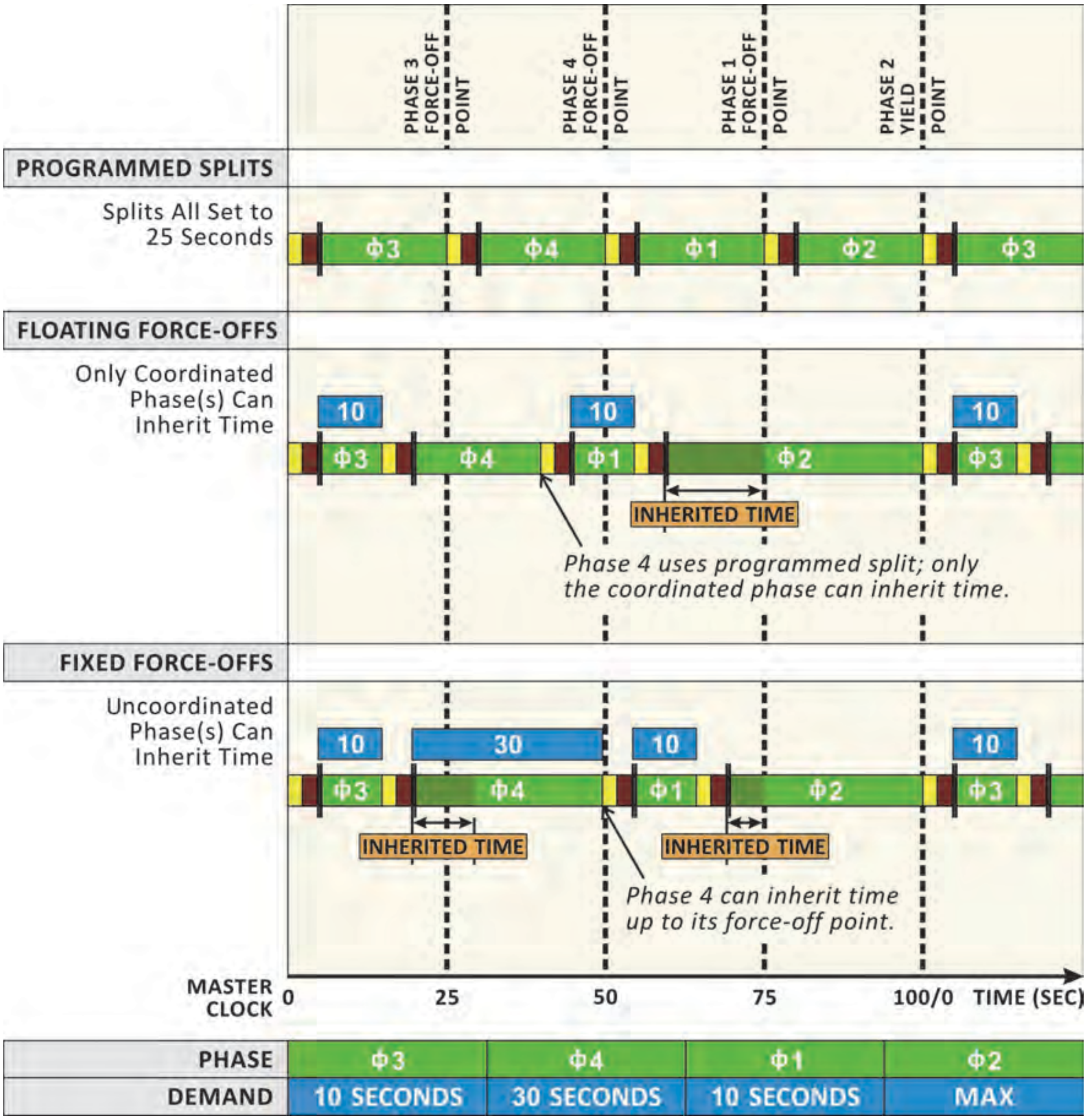


Exhibit 7-14 Fixed and Floating Force-Offs

7.3.5 Permissives

Permissives are windows of time when the controller can leave the coordinated phase(s) to serve the uncoordinated phases. In other words, they define the earliest and latest points at which an uncoordinated phase can ***begin*** in each ring. When a permissive period is active, the first uncoordinated phase in the sequence with demand will begin (i.e., an earlier phase in the sequence without demand will be skipped). Once the controller yields the coordinated phase to an uncoordinated phase, all subsequent phases with calls will be served; permissives no longer have an effect on the current

Permissives tend to be firmware specific.

cycle once the coordinated phase has yielded. Permissive periods are influenced by both the walk mode (discussed in detail in Section 7.5.1) and the type of permissive period that is programmed.

A practitioner will not be able to define permissive periods in every controller; most controllers automatically calculate and apply the largest permissive period possible (based on the vendor-specific approach to permissives). Regardless of whether permissives are defined by the practitioner or the controller, there are two general methods by which they are applied: (1) ***simultaneous*** permissive periods that start all permissives at the same time and (2) ***sequential*** permissive periods that start one at a time (i.e., the second permissive opens when the first closes, the third permissive opens when the second closes). The advantage of simultaneous permissives, under light traffic volumes, is the ability to skip Phase 4 and call Phase 1 early (normal sequence being Phases 1, 2, 3 and 4) if the Phase 1 call comes before Phase 4. Understanding the differences between these two methods will help a practitioner understand how a controller works under very low traffic volumes.

7.3.6 Yield Point

The yield point is the beginning of the first (or all) permissive periods and essentially determines the earliest point at which the coordinated phase(s) can begin termination and transition to uncoordinated phases. A yield point exists for each coordinated phase in each ring for most controllers, and while its primary purpose is to define a point in time when the controller can begin to terminate the coordinated phase(s), it is also the point at which transitions to new time-of-day plans typically take place. If lead-lag left-turn phasing is used for the movements adjacent to the coordinated phases, care should be taken to understand sync reference (as described below).

7.3.7 Pattern Sync Reference

Pattern sync reference is the start of the master clock (as outlined in NTCIP 1202, *2*) that occurs daily. As defined previously, the master clock is the background timing mechanism within the controller logic to which each local controller is referenced for time-based coordination. Pattern sync reference may occur at midnight (as shown in Exhibit 7-1) or other times during the early morning (e.g., 1:00 a.m. or 3:00 a.m.) as selected by the practitioner or defined by the firmware. A reliable time reference must be established (e.g., GPS-based time reference) because local controller clocks must keep the same time for coordination to be most effective (i.e., the clocks must be synchronized).

7.3.8 Offset Reference Point

Offset reference points help structure the relationship among coordinated intersections by defining the point in time (relative to the master clock) when the cycle begins timing. Reference points can be defined using a variety of points in a cycle, as shown in Exhibit 7-15 (which assumes Phases 2 and 6 are the coordinated phases). Many modern controllers allow several options including

- Beginning of ***first*** coordinated phase green,
- Beginning of ***both*** coordinated phase green,

- Beginning of ***first*** coordinated phase flashing don't walk (FDW), and
- Beginning of ***first*** coordinated phase yellow (end of first coordinated phase green).

The distinctions of "first" and "both" for the green, yellow, and FDW indications are important if the coordinated phases are using different values or the sequence has lead-lag left turns. While any offset reference can be computed from another offset reference (using cycle length and split times), the use of different offset references in the same system will produce poor results. This manual uses the beginning of first coordinated phase yellow (i.e., end of first coordinated phase green) as the offset reference because it is generally the most readily observable in the field.

Offset reference points should be used consistently.

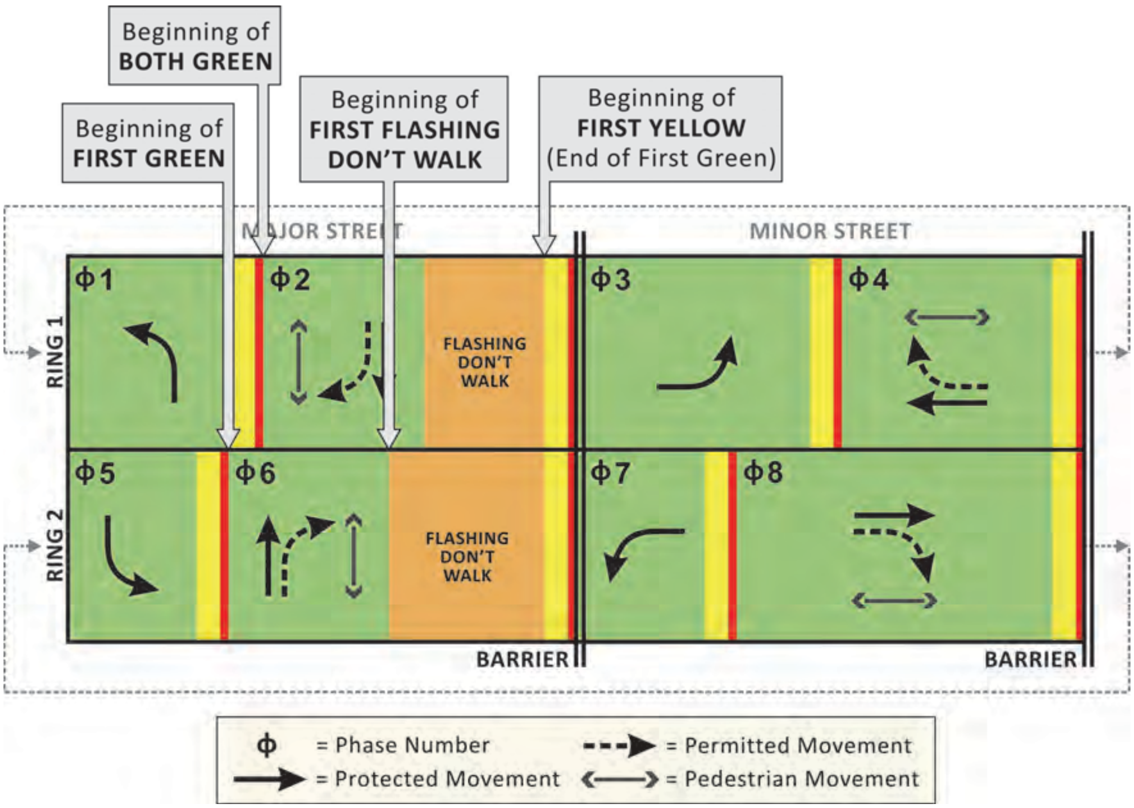


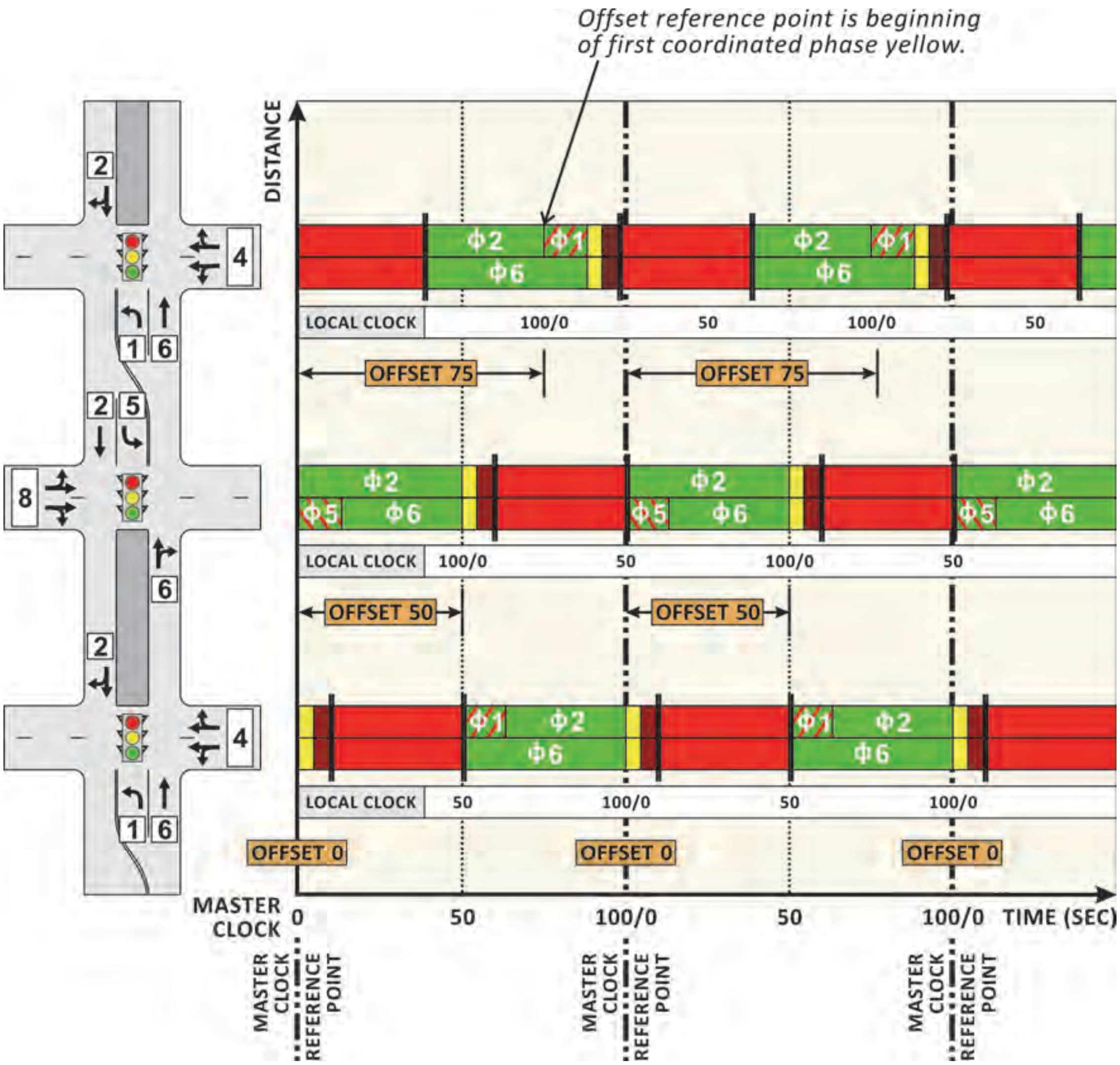
Exhibit 7-15 Offset Reference Points

7.3.9 Offsets

Offsets define the time relationship (expressed in seconds or as a percentage of the cycle length) between the master clock and the coordinated phases at local intersections. Once an offset reference point has been identified, an offset is defined as the time that elapses between the beginning of the master clock cycle (master zero point) and the offset reference point at the local intersection (local zero point). Intersection offsets of 0 seconds, 50 seconds, and 75 seconds from the master clock are shown in Exhibit 7-16.

In Exhibit 7-16 (and throughout this manual), the offset reference point is the beginning of first coordinated phase yellow. Note the lagging left-turn sequence at the northern-most intersection. Because the offset reference point is the beginning of first coordinated phase yellow, the offset reference point occurs at the end of Phase 2. In this example, the practitioner would see the Phase 2 yellow interval at the northern-most intersection begin 25 seconds after the Phase 2 (and Phase 6) yellow interval at the middle intersection.

Exhibit 7-16 Offsets



7.4 COORDINATION PARAMETER GUIDANCE

Once a practitioner has an understanding of the interactions and influence of coordinated timing parameters, he or she can use coordination planning tools (such as time-space diagrams) to select values for the coordinated parameters and create a timing plan. The following section provides additional guidance for this process.

7.4.1 Coordinated Phases Guidance

The desired outcome of coordination is often progressing the heaviest movements along a corridor, which frequently results in the major street through phases acting as the coordinated phases. (As explained in Chapter 5, these movements are typically designated as Phases 2 and 6.) However, the practitioner should explicitly consider traveler origin-destinations and traffic volumes when selecting the coordinated phase(s) because even a phase associated with a turning movement can be assigned as the coordinated phase.

7.4.2 Cycle Length Guidance

Cycle lengths, in combination with splits and offsets, establish relationships between intersections. Some cycle lengths will work better than others due to the time-space relationship between intersections, particularly if there is a desired system outcome (e.g., progression speed). Cycle lengths are frequently selected to address operations at a critical (or highest volume) intersection in a group of coordinated signals, but several additional cycle length selection methods are discussed in this section.

Cycle length selection is typically based on traffic data that are collected during representative time periods. If available, automated data collection provides more robust information for selecting cycle lengths.

The practitioner should always consider whether intersections with long cycle length requirements, especially those with adequate separation from other intersections (i.e., no queue spillback to adjacent intersections), would better operate independently from a group. Cycle lengths that are too long may increase congestion, rather than reduce it, due to long wait times experienced by queued vehicles on both the minor and major streets. Longer cycle lengths tend to increase congestion when

- More vehicles move through an intersection than can be handled downstream.
- There are turn-lane storage bay issues. Long cycle lengths may cause vehicles in turn bays to back up into through lanes. In a similar manner, long cycles may cause through traffic to back up beyond turn-lane storage bays, restricting access for turning vehicles.
- Headways increase (reducing flow rate at the stop bar), as queued vehicles leave through lanes to enter turn lanes.
- There is increased variability in the actuated green times. Long cycle lengths can result in high variability in the green time used by the minor street, which may result in fewer vehicles arriving on green at the downstream intersection. This is particularly noteworthy when split times exceed 50 seconds (*3*).

7.4.2.1 Manual Methods for Cycle Length Selection

Although most signal timing plans are developed using computer-based analysis, the tools are only as good as the traffic data and the optimization model. Often, too much emphasis is placed on the tool and too little on understanding the basic relationships. In many simple networks with modest traffic volumes, timings can be developed manually by selecting an appropriate cycle length (and offsets) based on block spacing.

A manual method for determining cycle lengths in a traffic signal network was first developed in Los Angeles (the City). The City used a traffic signal timing strategy that was elegantly simple for its then master-planned robust grid system. Most of the traffic signals in the San Fernando Valley part of the City operated using just two phases and permitted left turns, with signals spaced every quarter-mile. A 60-second cycle length was selected, and the green time was equally distributed between the minor street and major street traffic. Given the timing parameters and intersection spacing, a vehicle traveling at 30 miles per hour (mph) would reach the next quarter-mile signal in 30 seconds, or half of a cycle length. This resulted in an “alternating” system of offsets between adjacent intersections (as shown in the time-space diagram in Exhibit 7-17). With 30-second splits, progression was achieved on both the major and minor streets.

Exhibit 7-17
Alternating System of Offsets

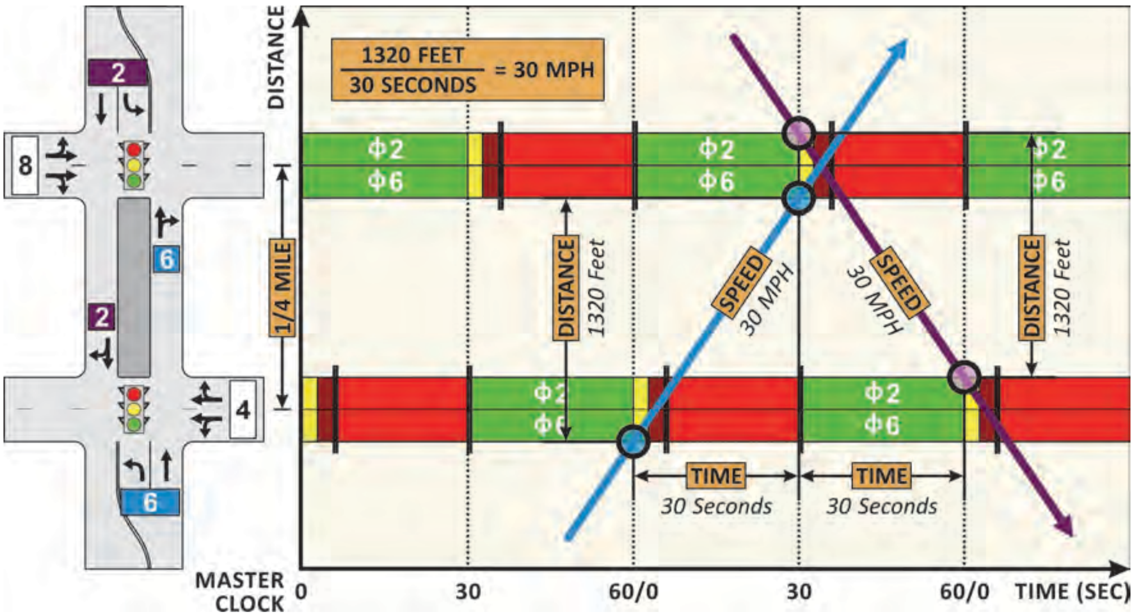


Exhibit 7-18 provides additional guidance on the relationship between cycle length, block spacing, and speed. The speed of progression can essentially be determined by dividing the block spacing by one-half, one-fourth, and one-sixth of the cycle length, respectively, for single, double, or triple alternate systems (depicted in Exhibit 7-19, Exhibit 7-20, and Exhibit 7-21).

Exhibit 7-18
Progression Speeds for Predetermined Cycle Lengths and Equal Block Spacing

Block Spacing (Feet)	Progression Speed (MPH)								
	60-Second Cycle Length			*90-Second Cycle Length*			*120-Second Cycle Length*		
	Single	*Double*	*Triple*	*Single*	*Double*	*Triple*	*Single*	*Double*	*Triple*
330	8	15	23	5	10	15	4	8	11
660	15	30	45	10	20	30	8	15	23
1320	30	60	90	20	40	60	15	30	45
2640	60	120	180	40	80	120	30	60	90

Note: Patterns that are grayed out are not recommended because of undesirable speeds.

In single alternate systems, every other intersection has an offset equal to half the cycle length relative to the master clock. In double alternate systems, every pair of intersections has a half-cycle offset relationship. In triple alternate systems, every three adjacent intersections have a half-cycle offset with the master clock. While triple alternate systems have relatively poor progression bands, the simultaneous offsets may be effective with high turning volumes from the minor streets. Minor street traffic would get an early green (explained in Section 7.6.2) at the middle intersection, potentially clearing queues before through traffic arrived.

Depending on the block spacing, desired progression speed, and type of alternate system, the practitioner should be able to determine an appropriate cycle length using the information in Exhibit 7-18. For example, if a practitioner wants to time signals along a corridor with quarter-mile spacing, using a single alternate system, for a desired progression speed of 30 mph, the signals should be timed using a 60-second cycle length.

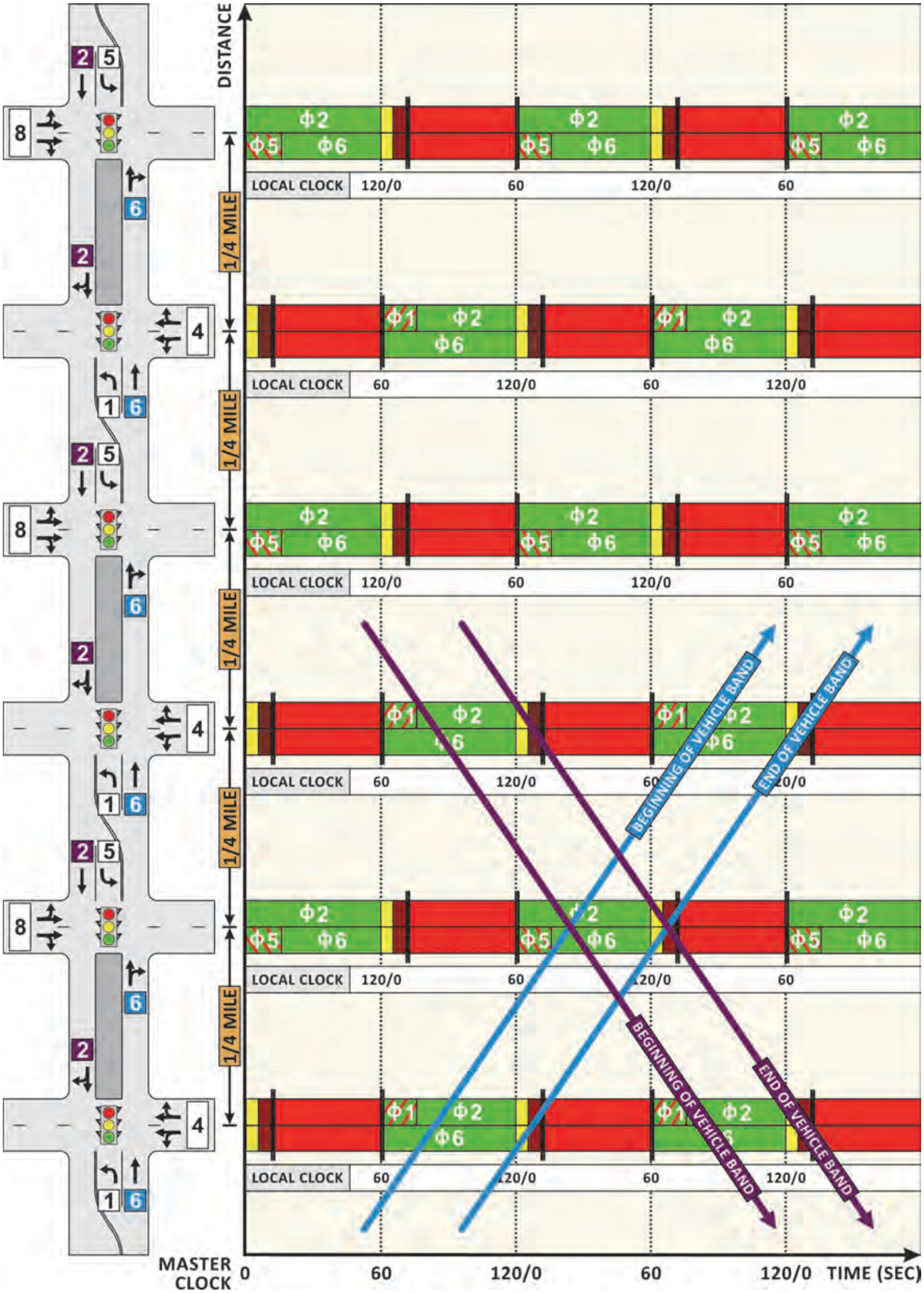
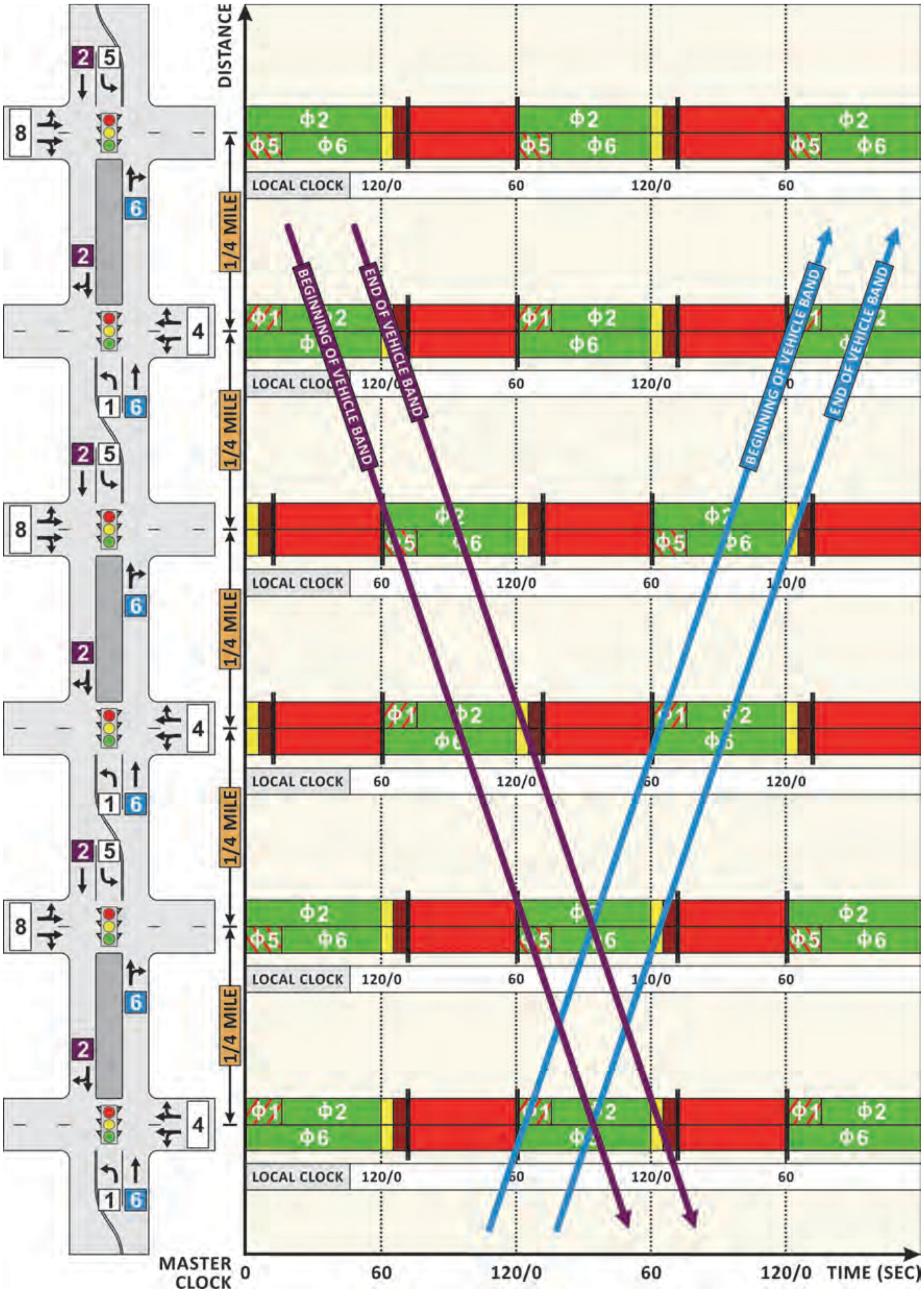


Exhibit 7-19 Single Alternate Coordinated System

Exhibit 7-20 Double Alternate Coordinated System



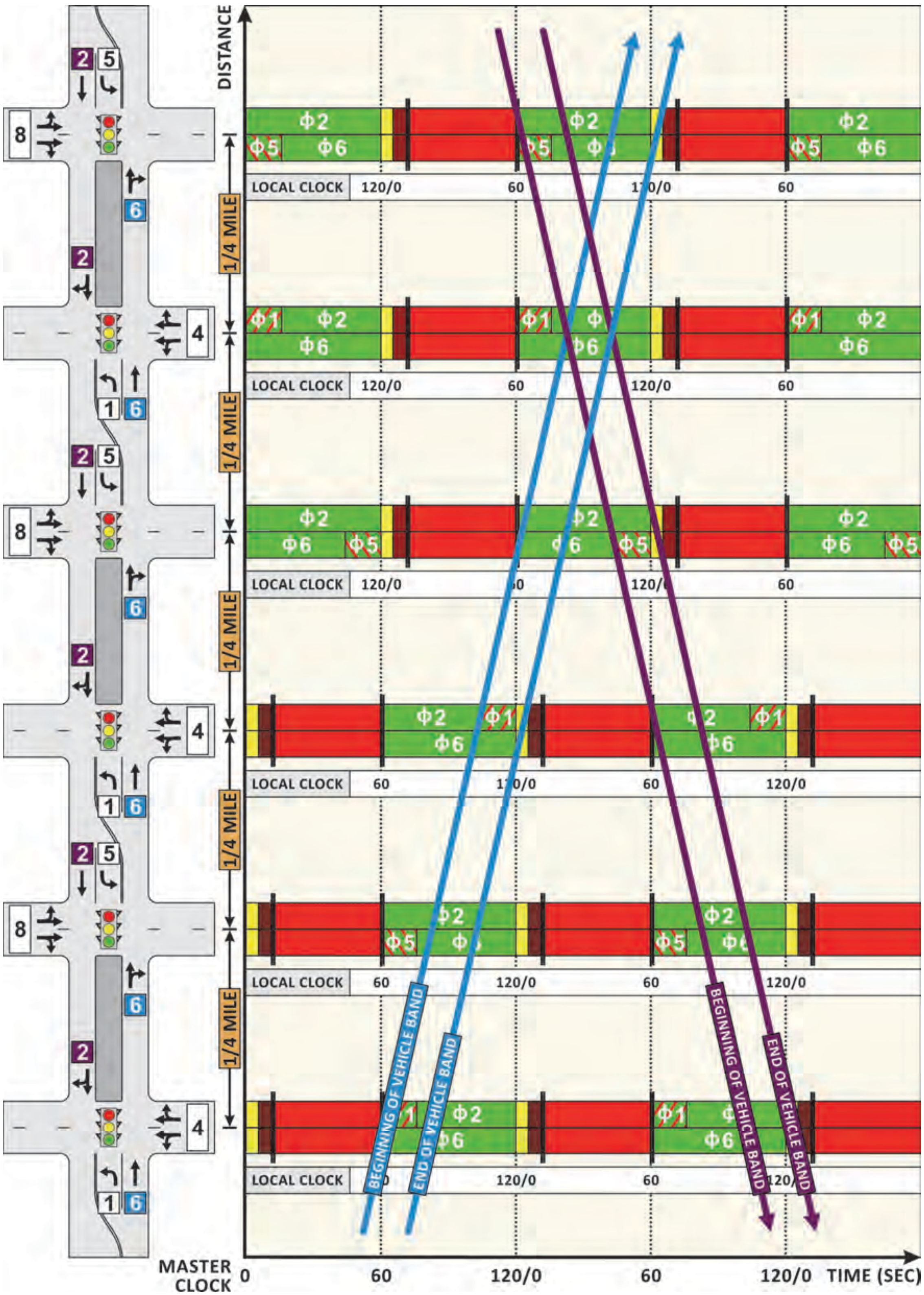
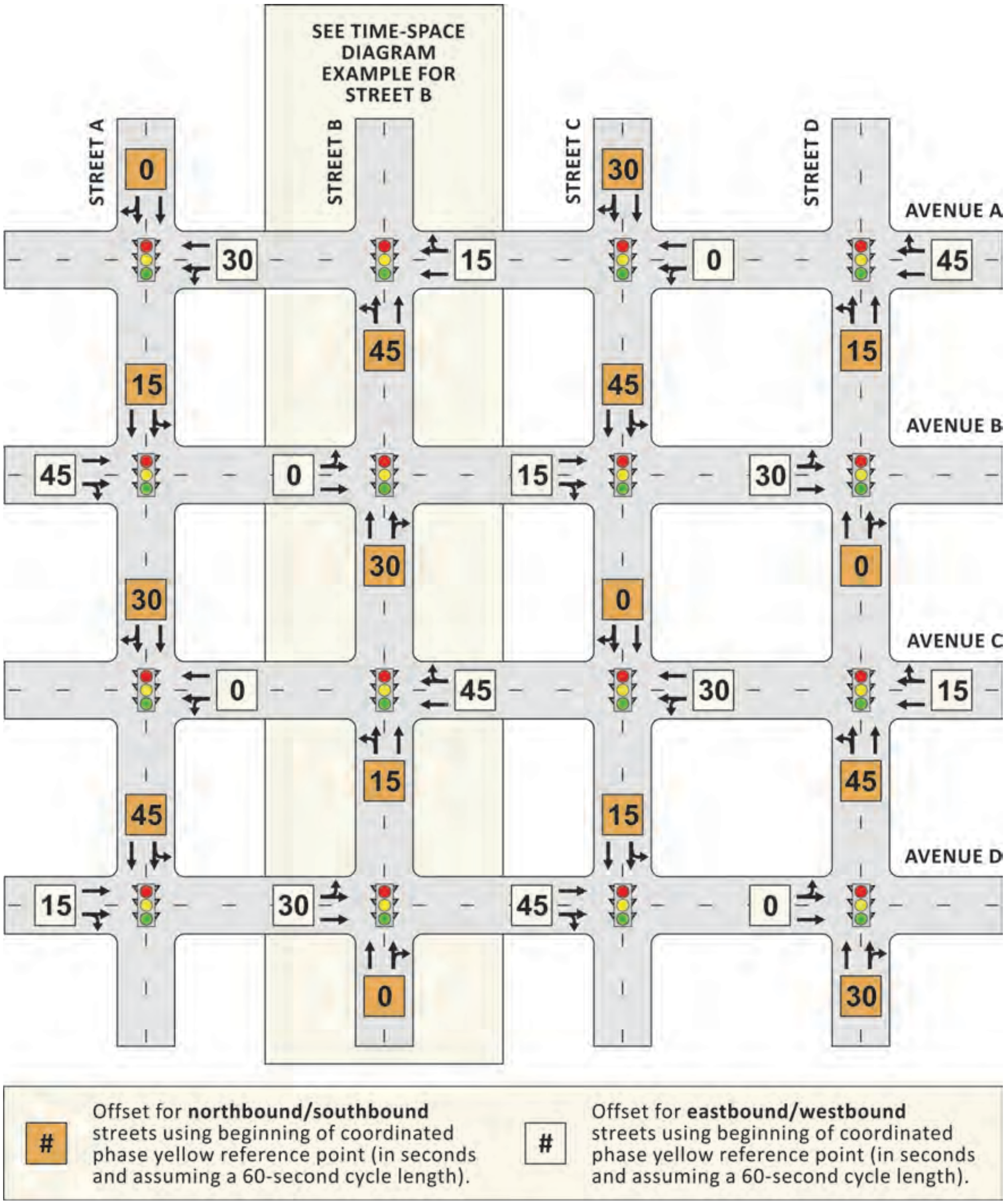


Exhibit 7-21 Triple Alternate Coordinated System

Manual methods for developing coordinated signal timing can also be applied to downtown grid networks. This method has been deployed in downtown Portland, Oregon, by separating intersections into a quarter-cycle offset pattern (illustrated in Exhibit 7-22 and Exhibit 7-23). The block spacing in downtown Portland is fairly uniform and relatively short (280 feet), and the grid is a one-way network (resulting in each intersection operating with two phases). Each subsequent intersection is offset by a quarter of the cycle length, which was selected to progress traffic on both streets. The result is a 13 mph progression speed and a 60-second cycle length. Note that if a cycle length is short, pedestrian coordination can also be achieved in the opposite direction of the one-way vehicular movements.

Exhibit 7-22 Example of Quarter-Cycle Offset Coordinated System



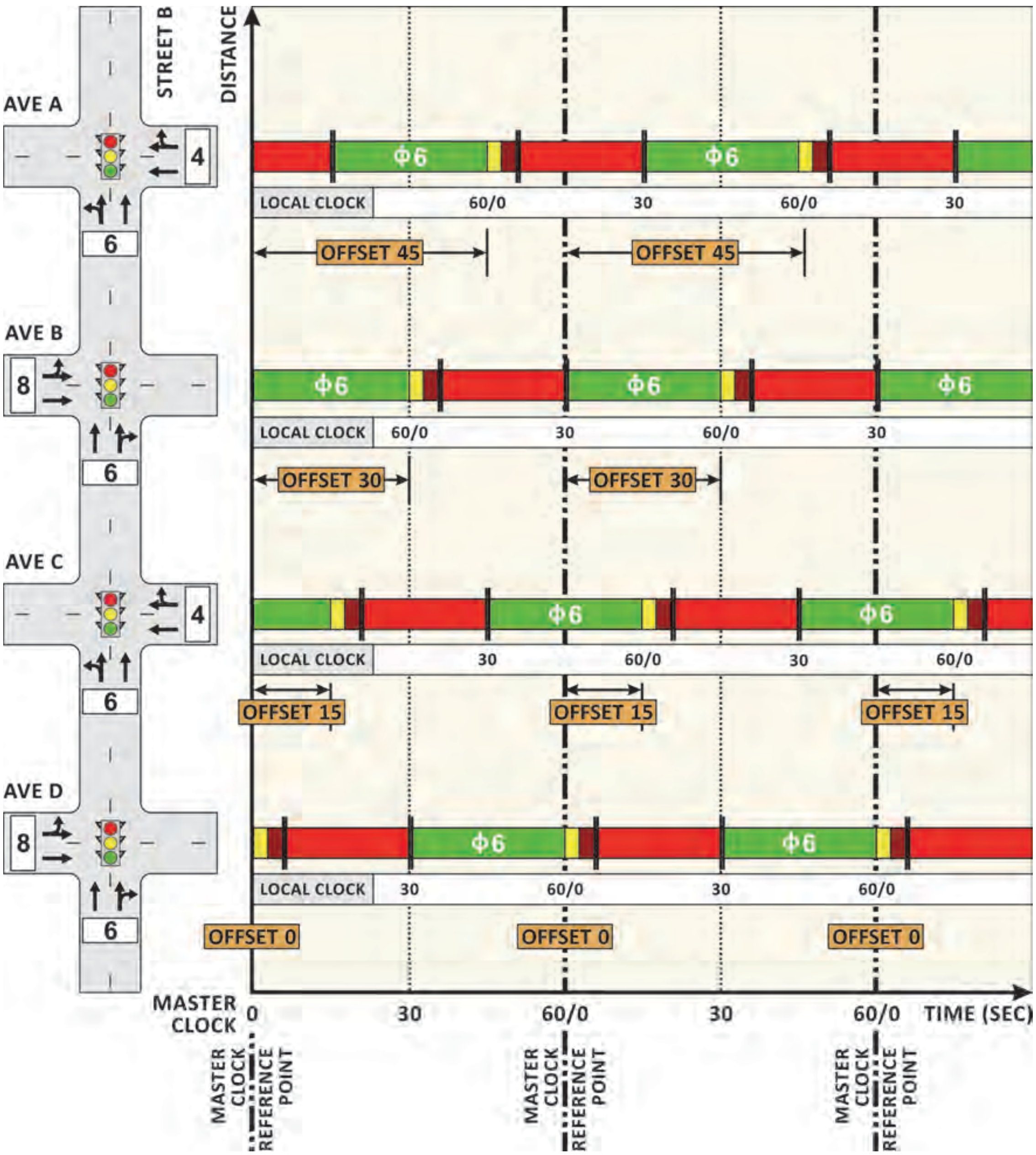


Exhibit 7-23 Example of Time-Space Diagram for Quarter-Cycle Offset Coordinated System

7.4.2.2 Critical Intersection Methods for Cycle Length Selection

As mentioned previously, the cycle length for a coordinated group of intersections can be based on the cycle length required at the critical intersection. Using this methodology, a cycle length is established that will sufficiently maintain under-saturated conditions at the critical intersection. While there are several critical intersection methods, the traditional method uses Webster's model (*4*) to determine optimal cycle length. The formula is as follows:

The critical intersection is typically the intersection with the highest demand.

Equation 7-1

$$C = \frac{1.5L + 5}{1.0 - Y}$$

where

C = optimum cycle length (seconds),

Y = critical lane volume divided by the saturation flow, summed over the phases, and

L = lost time per cycle (seconds).

The critical intersection approach considers signalized intersections in isolation, and for that reason, may not always yield the optimal cycle length for system progression (if that is the desired outcome). This method does not consider intersection constraints beyond the lost time and saturation flow rate and has a *de facto* outcome focused on vehicular flow. It does not specifically consider the needs of other users, such as pedestrians, bicycles, and transit, and is only valid for under-saturated conditions. (Oversaturated conditions require special consideration; more information can be found in Chapter 12.)

7.4.2.3 Network Approach to Cycle Length Selection

The network approach to cycle length selection considers the performance of multiple intersections when determining an optimal cycle length. Most applications use signal timing optimization models that are based on vehicular flow. The FHWA *Signal Timing Process Final Report* and *Traffic Analysis Toolbox* describe several available optimization software packages (*5, 6*). However, optimization models change with new versions of software, and, therefore, documentation is often best obtained from individual software producers.

Optimization models consider the network being analyzed and determine an optimal solution based on a given set of inputs (including the range of preferred cycle lengths). These models generally use individual intersection characteristics, volume-to-capacity ratios, link speeds, and the distance between intersections to estimate the performance of individual cycle lengths and resulting plans. The timing policies, optimization policies, and criteria used to select a signal timing plan must be considered prior to and as a part of the cycle length optimization process. A significant amount of effort may be required to take an initially screened plan from an optimization model to a point that can be field implemented.

7.4.3 Splits Guidance

In many controllers, maximum green values may be ignored during coordination using the Inhibit Max feature. This allows a phase to use a split time that is longer than its normal maximum green value.

Split distribution is the process of determining how much of the cycle can be provided to each of the phases. Determining adequate split times can be challenging. If a split time is too long, other phases may experience increased delay (if detection and associated passage time do not allow for termination at the end of maximum flow). If a split time is too short, the demand may not be served.

There are various ways to determine the necessary split time for a phase. The general intent is to allocate enough time to each phase to avoid oversaturated conditions for consecutive cycles, but over the course of an analysis period (15 minutes or 1 hour), splits should be distributed in a manner that is consistent with the operational objectives. For example, if the operational objective is minimizing stops on a corridor, the corridor phases should likely receive longer split times relative to the minor street phases.

One method for split distribution is allocating splits that satisfy a design level of capacity for all of the minor movements, with the remaining residual time allocated to the coordinated movements. Another common practice allocates the green time such that the volume-to-capacity ratios for critical movements are equal. With the many interactions between signal timing parameters, the allocation of green time should also depend on pedestrians, transit phases, and bandwidth considerations.

7.4.4 Force-Offs Guidance

Controllers often automatically calculate force-offs based on the programmed splits. However, some controllers may allow the practitioner to program the force-off values, or the type of force-off that is applied. The phase directly after the coordinated phase will not have an opportunity to receive time from a preceding phase, regardless of the method of force-off (unless the coordinated phase is actuated). Thus, the force-off (or split) for the first phase following the coordinated phase should be selected carefully.

Depending on the operational objectives, consideration should also be given to the type of force-off that is applied. Floating force-offs favor the coordinated phase(s), as they do not allow uncoordinated phases to inherit time. Fixed force-offs, on the other hand, can be beneficial if there are fluctuations in traffic demand, and an uncoordinated phase needs more green time during a cycle. This type of force-off can also help to prevent early return to green on the coordinated phases (*7*), reducing perceived delay along the corridor (described in detail in Section 7.6.2).

One of the potential outcomes of fixed force-offs is that a phase later in the sequence (before the coordinated phase) may receive more than its split time (provided the maximum green timer is either not active or is not reached). Inhibit Max (a definable controller parameter) may be invoked to prevent the controller from using maximum green values during coordinated operations. If allowed to function during coordination, maximum green values could result in phases not reaching their force-off points.

Inhibit Max must only operate during coordinated time periods.

7.4.5 Permissives Guidance

Many modern controllers automatically maximize permissives, but they do not all operate in the same manner, as discussed earlier. Larger permissive periods are desirable during times of day with low traffic volumes because of the increased opportunities for uncoordinated phases to be served. With higher traffic volumes, permissives are not a significant issue because a call usually exists on all phases when the controller reaches the yield point. Maximizing permissives is generally recommended.

7.4.6 Yield Point Guidance

Yield points are ultimately determined based on when permissive periods begin. Modern controllers calculate yield points based on the offset reference and walk mode. No other consideration is necessary.

7.4.7 Pattern Sync Reference Guidance

Pattern sync reference should generally be programmed at a time when there are lower traffic volumes in order to limit disruptions to traffic flow. It is important that each intersection reference a consistent master clock so that local controllers use identical reference points. Each controller should be configured to keep track of subtle time-related issues that will keep it in sync with the master clock, such as if the area follows daylight savings time and/or when daylight savings time begins and ends. If the local clock time is not synchronized with the master clock, coordination will not function correctly.

7.4.8 Offset Reference Point Guidance

Offset reference points will not move based on when phases actually time; they reference a consistent point in time that is based on phase splits and sequence. Reference points are generally programmed through controller or system firmware, and it is essential for the practitioner to apply a consistent offset reference point for a group of coordinated signals. However, it is important to realize that all offset reference points are equivalent operationally (i.e., if you know one reference, you can calculate any other reference for the same pattern).

The beginning of first coordinated phase green has been a common offset reference point and is the official value defined in NTCIP 1202 (*2*). However, it is not readily observable in the field because uncoordinated phases can terminate early through gap outs. This causes the coordinated phases to begin earlier than the programmed "beginning of first coordinated phase green" offset reference point.

The examples throughout this manual use the beginning of first coordinated phase yellow as the offset reference point.

The beginning of first coordinated phase yellow is generally observable or apparent in the field (i.e., the actual point does not move as a result of uncoordinated phases timing less than their allocated splits). One exception to an observable beginning of yellow is if the coordinated phases are actuated. In that case, only the beginning of FDW is observable (although not a typical offset reference).

7.4.9 Offsets Guidance

Offsets should be chosen based on the actual or desired travel speed between intersections, distance between signalized intersections, and traffic volumes. In an ideal coordinated system, offsets would allow platoons (leaving an upstream intersection at the start of green) to arrive at a downstream intersection near the start of green or after the queue from minor streets or driveways is discharged (i.e., green starts early enough to clear queued vehicles before platoon arrives).

Offsets are often adjusted in the field, but field observations provide the practitioner with a limited ability to review conditions. With a 100-second cycle, there are only 36 cycles during an hour, and one cycle may not be indicative of the next cycle's performance. Instead of attempting to base offsets solely off of field observations ***or*** software output, the practitioner should use field review ***and*** time-space diagrams in combination to optimize the system. Some modern controllers also monitor arrivals on green as an aid to fine-tuning offsets.

7.5 OTHER CONSIDERATIONS FOR COORDINATION

Beyond basic coordination parameters, there are other considerations for coordinated operations. A practitioner should understand how walk modes, actuating the coordinated phase(s), and transition logic will affect operational objectives.

7.5.1 Pedestrian Timing and Walk Modes

Pedestrian operations can have a direct impact on the coordination along a corridor. Pedestrian timing is required for all phases that serve pedestrians. However, when pedestrian service is actuated and demand is relatively low, it may be desirable to allocate a split time that is shorter than the time required to serve a pedestrian (if the controller firmware supports the capability). Traffic operations may be more efficient without accommodating pedestrians within the coordinated cycle length (if the pedestrian time is longer than what is needed for vehicular progression and traffic demand). It may be more effective for a controller to be shifted out of coordination and have to transition back for the occasional pedestrian than to serve pedestrian timing every cycle.

7.5.1.1 Pedestrian Timing for Uncoordinated Phases

The time necessary to walk across the street (pedestrian timing requirements) may be longer than the needs of other traffic. The effect of pedestrian timing on coordination is, therefore, most apparent as it affects the timing on the minor street. Exhibit 7-24 illustrates the basic principle of pedestrian timing when the vehicle split is sufficient to accommodate the required pedestrian time, allowing the signal to stay in coordination (and not exceed the force-off point).

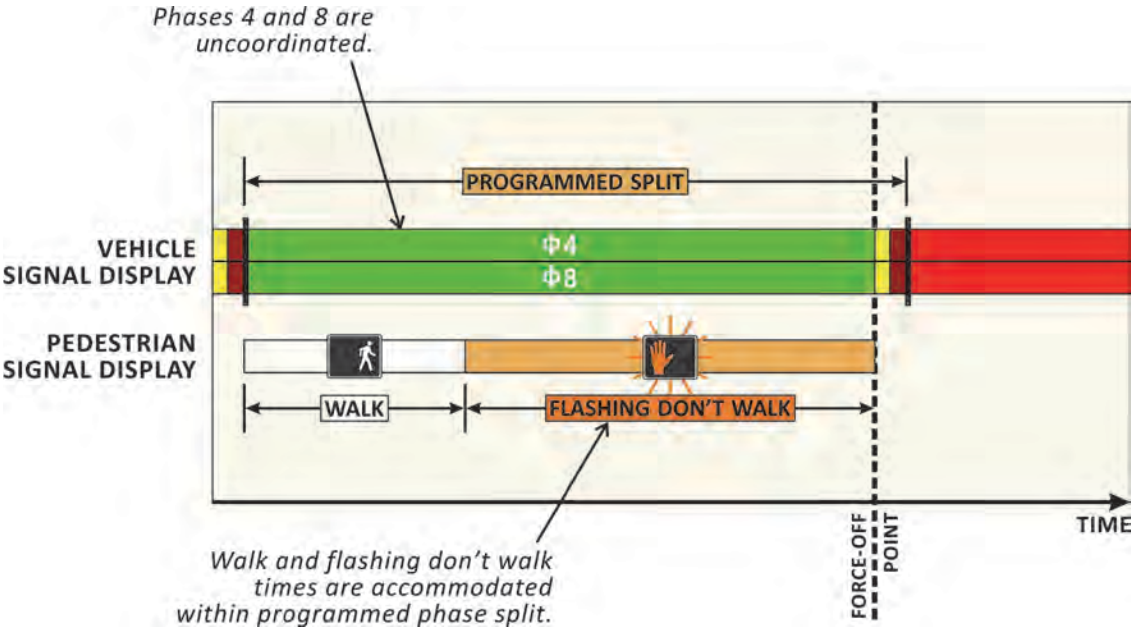
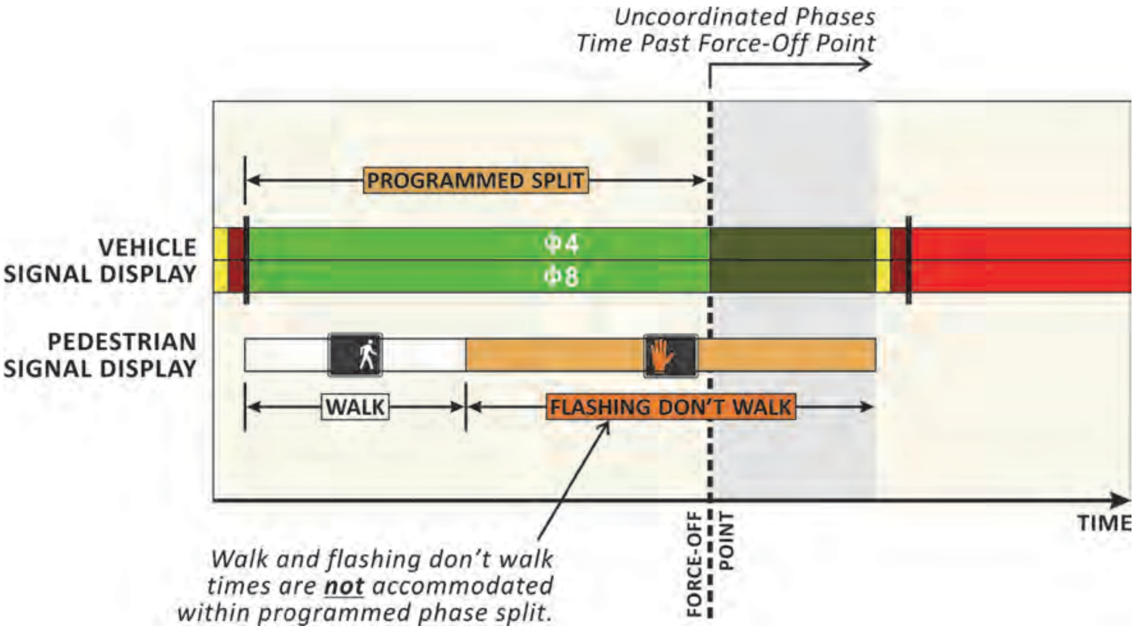


Exhibit 7-24
Uncoordinated Phase Operation with Pedestrian Timing Completed Before the Force-Off

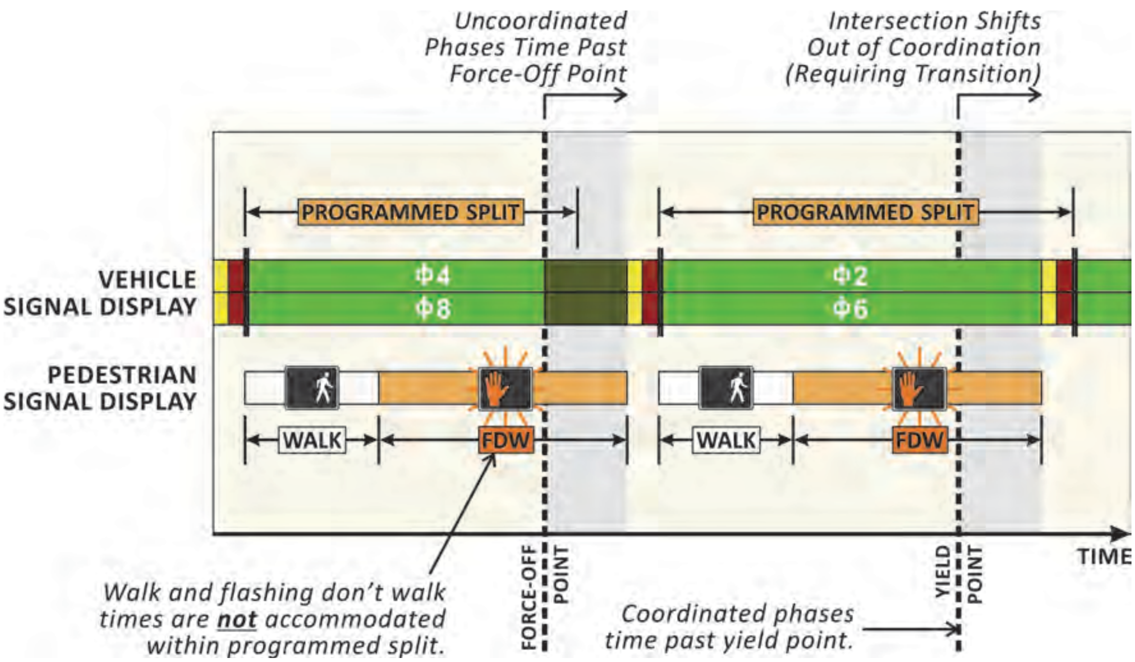
When the split for the subject phase is not sufficient to cover the pedestrian timing, the controller times the phase beyond its force-off point (as illustrated in Exhibit 7-25). The response of the controller depends on two factors: (1) demand for subsequent uncoordinated phases and (2) non-actuated versus actuated operations on the coordinated phases.

Exhibit 7-25 Uncoordinated Phase Operation with Pedestrian Timing Exceeding Phase Split



When the coordinated phases are non-actuated (which is typical practice), the coordinated phases must begin timing sufficiently in advance of the controller's yield point to enable full vehicle timing (minimum green) and pedestrian timing (walk plus FDW). Should the amount of time be insufficient to cover these timing requirements, the controller will time the coordinated phase past the yield point and fall out of coordination (as shown in Exhibit 7-26).

Exhibit 7-26 Loss of Coordination Due to Pedestrian Call



Signal Timing Manual, Second Edition

It is at the yield point that the controller logic typically determines the method by which the controller will transition back into coordination. If the late return to the coordinated phase is minor (relative to cycle length and requirements for other phases), it could provide minimal disruption and significant benefits to overall traffic, provided the transition mode allows the controller to take time from following phases (see Section 7.5.3).

7.5.1.2 Pedestrian Timing for Coordinated Phases

The amount of time needed to serve vehicular volumes or provide bandwidth along the major street usually results in coordinated phase splits that are sufficient to accommodate pedestrian timing. While progression considerations generally provide adequate time to accommodate pedestrians, some walk modes are more pedestrian-friendly than others, as discussed below.

7.5.1.3 Walk Modes

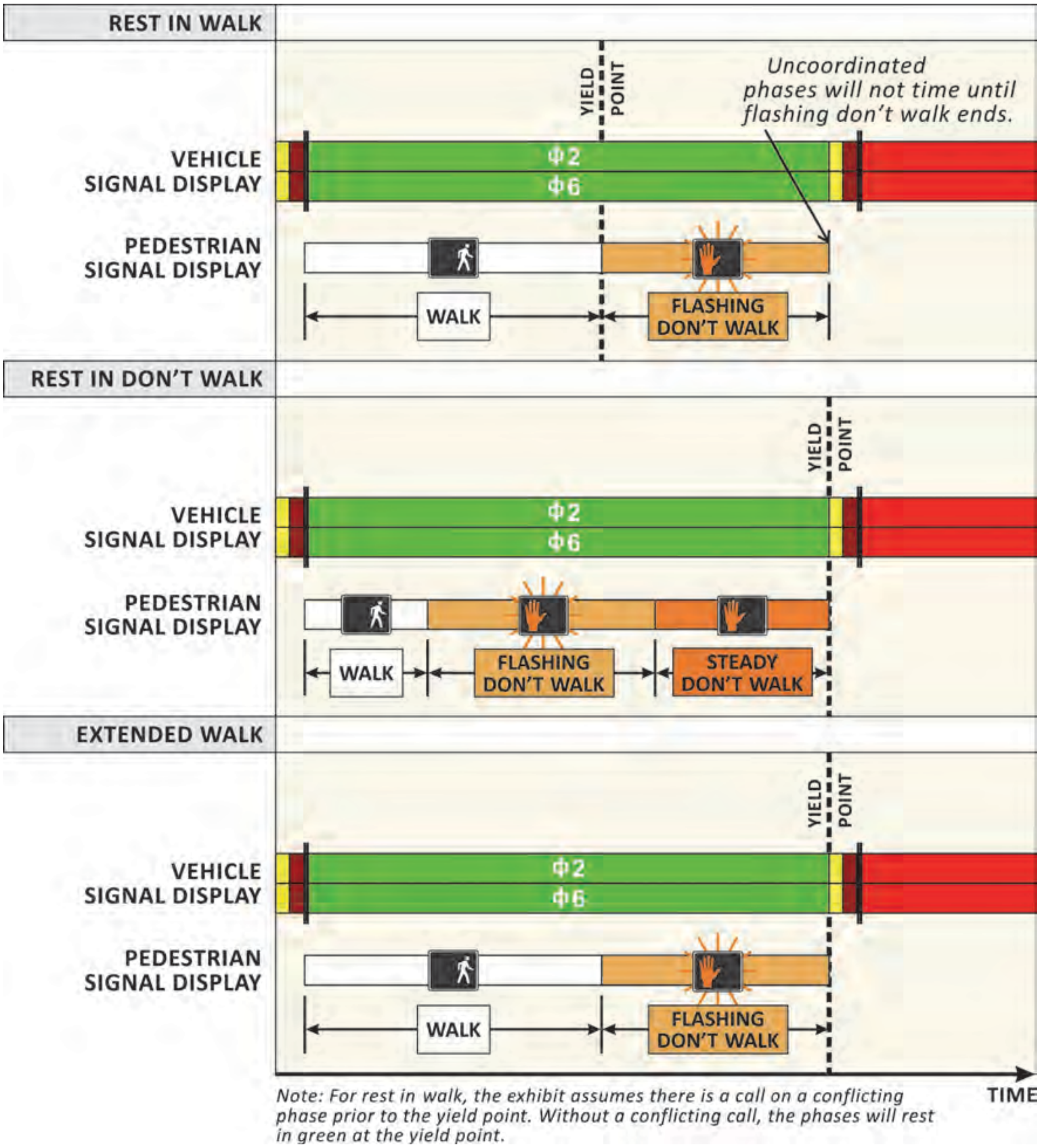
A practitioner can specify walk modes that influence how pedestrians are served during the coordinated phases: (1) rest in walk, (2) rest in don't walk, and (3) extended walk (Exhibit 7-27 illustrates these three walk modes):

- ***Rest in Walk*** dwells in the pedestrian walk interval while the coordinated phase is green, regardless of pedestrian calls. This mode is often used when there are high pedestrian volumes, such as in downtown environments or locations near schools, and it does not ***require*** any pedestrian detection (although pedestrian detection may be desirable to allow for late-night free operation). However, this walk mode causes the FDW interval to extend past the yield point, delaying minor street movements until the FDW interval has ended. The delay to minor streets is only noticeable under low-volume conditions.
- ***Rest in Don't Walk*** dwells in the steady don't walk interval after the programmed walk and FDW intervals have been served. This mode is often used when pedestrian volumes are low. It does require pedestrian detection.
- ***Extended Walk*** dwells in the pedestrian walk interval (similar to rest in walk) starting at the beginning of the coordinated phase green. It maximizes the walk time every cycle, but also times the FDW interval prior to the yield point. This mode is a compromise between the first two walk modes, and it does not require pedestrian detection.

The appropriate walk mode may depend on the time of day. During the middle of the night, rest in don't walk might be the most appropriate because of the low volume of pedestrians. However, a high volume of pedestrians during the day may make rest in walk (or extended walk) the most appropriate for that time period.

Depending on the walk mode and pedestrian actuation settings, the walk interval may be able to time more than once during a cycle. Pedestrian re-service is a feature that can be used when pedestrian service is actuated on the coordinated phase. It allows the walk interval to time again if (a) a call is placed and (b) there is enough time before the latest point at which the FDW interval must begin.

Exhibit 7-27
Pedestrian Walk Modes



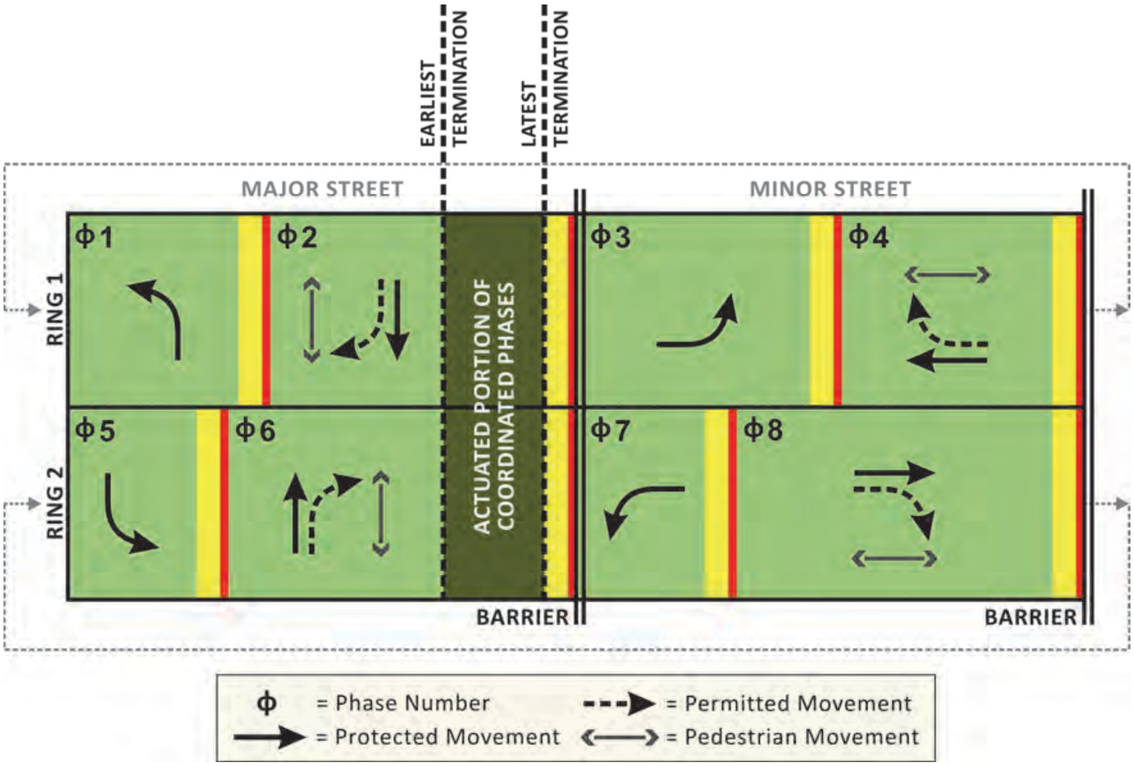
7.5.2 Actuating the Coordinated Phase

A portion of the coordinated phase can be actuated, which increases the likelihood of decision zone protection and offers more flexibility to serve minor street and left-turn movements.

A portion of the split time at the end of the coordinated phases can be actuated, allowing for an earlier termination of the coordinated phases in the absence of detected traffic demand (as illustrated in Exhibit 7-28, assuming rest in don't walk). This approach effectively moves the phase termination point earlier in the cycle and allows the coordinated phases to gap out and give time to the minor streets and left turns. In addition, it increases the likelihood of decision zone protection (through the use of detection to end the coordinated phases instead of a max out or force-off).

An actuation value that is just right will (1) allow for occasional gap outs to serve minor street demand more quickly (reducing intersection delay) and (2) reduce max outs or force-offs for the coordinated phases (enhancing decision zone protection). An "actuated time" value that is too large can result in more early returns to green and/or the coordinated movements gapping out just prior to the arrival of the platoon, resulting in poor signal progression, increased stops, and delay along the corridor. A value that is too small will have an inconsequential result. Actuating the coordinated phase requires engineering judgment and observation of signal operations during multiple times of day or days of the week to set reasonable actuation values (typically per ring).

Exhibit 7-28 Actuating a Portion of the Coordinated Phases



7.5.3 Transition Logic

Transitioning is the process of either entering into a coordinated timing plan from "free" operations or changing between two plans. Transitioning may also be necessary after an event such as preemption or loss of coordination (possibly due to a pedestrian time that exceeds the allocated split, as discussed in Section 7.5.1). In general, traffic signals do not operate within the same pattern parameters and cycle lengths at all times. The pattern may change during the day for a number of reasons:

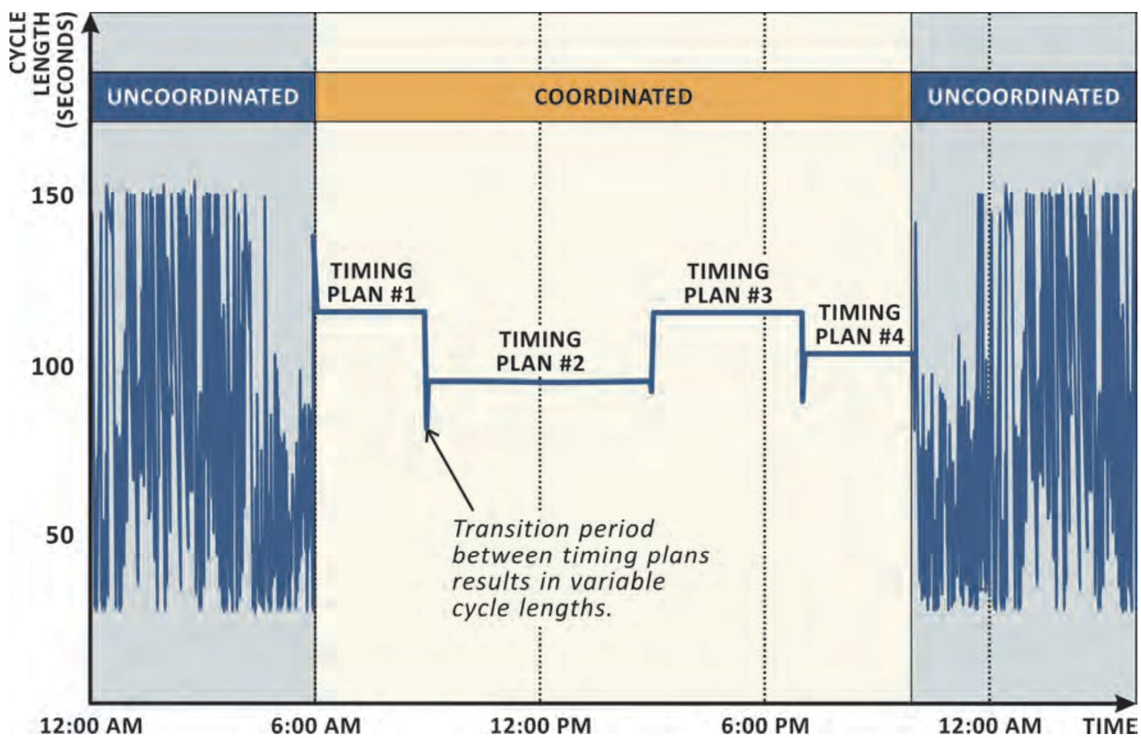
- Time-of-day scheduled changes;
- Manual operator selection;
- Traffic-responsive pattern selection;
- Emergency vehicle, railroad, or other preemption;
- Adaptive control system pattern selection;
- Corrections to the controller clock;
- Pedestrian time exceeds split time; or
- Power loss and restoration.

Studies have shown that excessive changes to timing plans, in an attempt to match traffic patterns closely and improve performance, can be a detriment because the system never achieves coordination long enough for the new plan benefits to outweigh the costs of transition. Because of this, it is generally recommended to remain in a coordinated pattern for at least 30 minutes. It is also best to avoid changing patterns during congested conditions when the signals need to operate at maximum efficiency. A peak-period pattern is best implemented early to ensure all offset transitioning is completed before the onset of peak traffic flows.

When the controller reaches a point when it is necessary to change the coordination pattern (the cycle, splits, offsets, and/or sequence), the controller may also have to shift the local offset reference point. This requires the use of an algorithm that may either shorten (i.e., subtract time from) or lengthen (i.e., add time to) the cycle. The transition algorithm typically operates over one to five cycles, depending on the transition mode selected and how much the offset reference point needs to shift. Consequently, the split durations (and cycle lengths) during the period of transition will be different from those programmed in either the previous pattern or new pattern.

For example, in the cycle length plot shown in Exhibit 7-29, the system runs with a fixed background cycle from 6:00 a.m. to 10:00 p.m., with cycle changes at 9:00 a.m., 3:00 p.m., and 7:00 p.m. Before 6:00 a.m. and after 10:00 p.m. the intersection is running under free (uncoordinated) operations, so the cycle length varies depending on the traffic detected at the intersection. During each of the plan changes, the controller goes into transition, resulting in variable cycle lengths for a few cycles to adjust to the new pattern. These variable cycle lengths are portrayed by the dips and rises in the overall steady cycle length line.

Exhibit 7-29 Example of Daily Cycle Length Fluctuations



While central- or master-based systems typically communicate the selection of a new timing plan to all signals in a group at the same time, the actual transition logic is typically executed independently at each signal, without explicit regard for the state of adjacent signals. Most controllers allow three or four transition modes, which govern the precise details of how the signal resynchronizes to the new offset reference point. The transition modes differ significantly from one controller manufacturer to the next. Some vendors may refer to transition as offset seeking, offset correction, or coordination correction (*8*). No matter which mode is selected, traffic control can be significantly less efficient during the transition between timing plans than it is during coordination.

The three most common techniques for achieving an offset reference point transition are

- Lengthening the cycle (adding time to each split),
- Shortening the cycle (subtracting time from each split), and
- Shortway (adding or subtracting time depending on which is faster).

To avoid a cycle length that is excessively long or green times that are too short during the transition period, it is common for controllers to limit the maximum amount of adjustment that can be made in one cycle. If such a limit is imposed (which is typical), the signal may not be able to complete a given transition within one cycle even if the adjustment is small. Signal controllers also commonly compute their adjustments so that transitioning is completed in a set number of cycles for the worst case scenario, typically a maximum of three to five cycles.

7.5.3.1 Lengthening Transition Modes

The most common lengthening transition modes provided by signal controllers in the United States include dwell, max dwell, and add only (which is generally preferable to dwell modes) (*9*). Lengthening transition modes offer less risk than transition modes that shorten the cycle length, unless the amount of shortening is known to be small. There is less likelihood of queues building up because of short green times and pedestrians not being served. However, if the offset reference point is being shifted 1 second backwards, using a lengthening transition mode requires shifting the entire cycle forward 1 second less than the cycle length. This results in longer cycles during the transition period, which could potentially cause unexpected storage problems in left-turn lanes or between closely spaced intersections. The modes are described as follows:

- ***Dwell:*** At the next display of green for the coordinated phases, the controller begins to transition by holding (or dwelling) in this state until the new offset reference point is achieved, at which time the signal is considered in sync and begins the new timing plan. This transition mode puts all the transition time into the coordinated phases, which may cause problems on other uncoordinated phases. If the offset reference point needs to move 1 second earlier, the coordinated phases will dwell for 1 second less than the cycle length.
- ***Max Dwell:*** This modified version of dwell also adjusts the start of the cycle by extending the green time of the coordinated phases. However, only a limited amount of extra green time may be added each cycle.

- ***Add:*** This mode synchronizes by shifting the start of the cycle progressively later, by timing slightly longer-than-programmed cycle lengths. The add mode increases the green time on all phases in the sequence, whereas the dwell modes add time only to the coordinated phases. If a signal is subject to preemption, selecting the add-only transition mode allows all phases to receive additional time during the transition period.

7.5.3.2 Shortening Transition Mode

The shortening transition mode shortens the cycle length, subtracting time from phases to the extent allowed by their minimum green settings and any pedestrian activity during the phase. Shortening can be effective if the offset correction is very small. Depending on the minimum green times and pedestrian times, there may only be small adjustments in cycle length that can be made by shortening. It may take many cycles to complete an offset reference point transition, so it is not practical to require a controller to use the shortening technique exclusively. To avoid this problem and allow use of the shortening transition mode when it works well, most controllers offer some version of the shortway method (discussed in the next section).

- ***Subtract:*** This mode shifts the start of the cycle progressively earlier, subtracting time from one or more phases in the sequence (subject to their minimum green time and pedestrian time requirements). This method is very effective for small corrections, such as pedestrian calls that go past the force-off by a small amount.

7.5.3.3 Shortway Transition Mode

Shortway is typically the least disruptive method that can be used to achieve an offset reference point transition.

If the practitioner selects the shortway transition mode, the controller will assess both lengthening and shortening techniques and automatically choose the one that will complete the offset reference point transition (i.e., get the signal "in step" or "in sync") most quickly. Field experience and laboratory experiments have shown that shortway typically provides the least disruptive effects on traffic compared to other methods, unless the signal is subject to preemption or the intersection is at near-saturated conditions. “Traffic Signal Transition in Coordinated Meshed Networks” reports that shortway provides quick and smooth transitions while avoiding severe peaks in delay. While this method was sometimes outperformed by others when a single approach was considered, it was always among the best methods in overall performance (*10*).

- ***Shortway:*** This mode (also sometimes called bestway, fastway, or smooth) finds the “shortest path” to transition the intersection by using either lengthening or shortening transition logic. The specific details of how this mode determines the “shortest path” can vary significantly from one controller vendor to the next.

7.6 COMPLEXITIES

This section discusses some of the various complexities of signal coordination. There are many variables that must be considered to achieve an acceptable coordination plan.

7.6.1 Phase Sequence

The sequence of phases, particularly left-turn phases, can significantly affect corridor operations. The most common phase sequencing decision—whether to lead or lag left turns—can have a particularly strong impact on bandwidth (in both directions) along a corridor. Other phase sequence decisions (such as the sequence of left turns on the minor street or the sequence of split phasing on the minor street) often have less impact on bandwidth and delay but should also be considered.

7.6.1.1 Major Street Left-Turn Phase Sequence

Lagging one of the major street left-turn phases (lead-lag) can facilitate better progression for both directions because it allows platoons to arrive at different times during the cycle (as demonstrated in Section 7.2.6). Depending on how the lagging left-turn phase is configured, it will generally receive the same amount of green time each cycle regardless of demand (unless the coordinated phase in the other ring has an actuated interval, discussed in Section 7.5.2). This fixed interval occurs because the coordinated phase in the other ring needs to end at the same time as the lagging left in order for both rings to terminate and cross the barrier together. Actuating the coordinated phase allows both the coordinated phase and lagging left to end early if demand is not present.

Modern controllers allow left-turn phase sequences to be varied by time of day. This has traditionally been done only for protected left-turn operations, but the use of FYA indications allows this to be extended to protected-permitted operations (see Chapter 4 for more information on FYAs). The practitioner should always consider user expectations and the operational objectives when choosing different phase sequences for different times of day.

7.6.1.2 Minor Street Left-Turn Phase Sequence

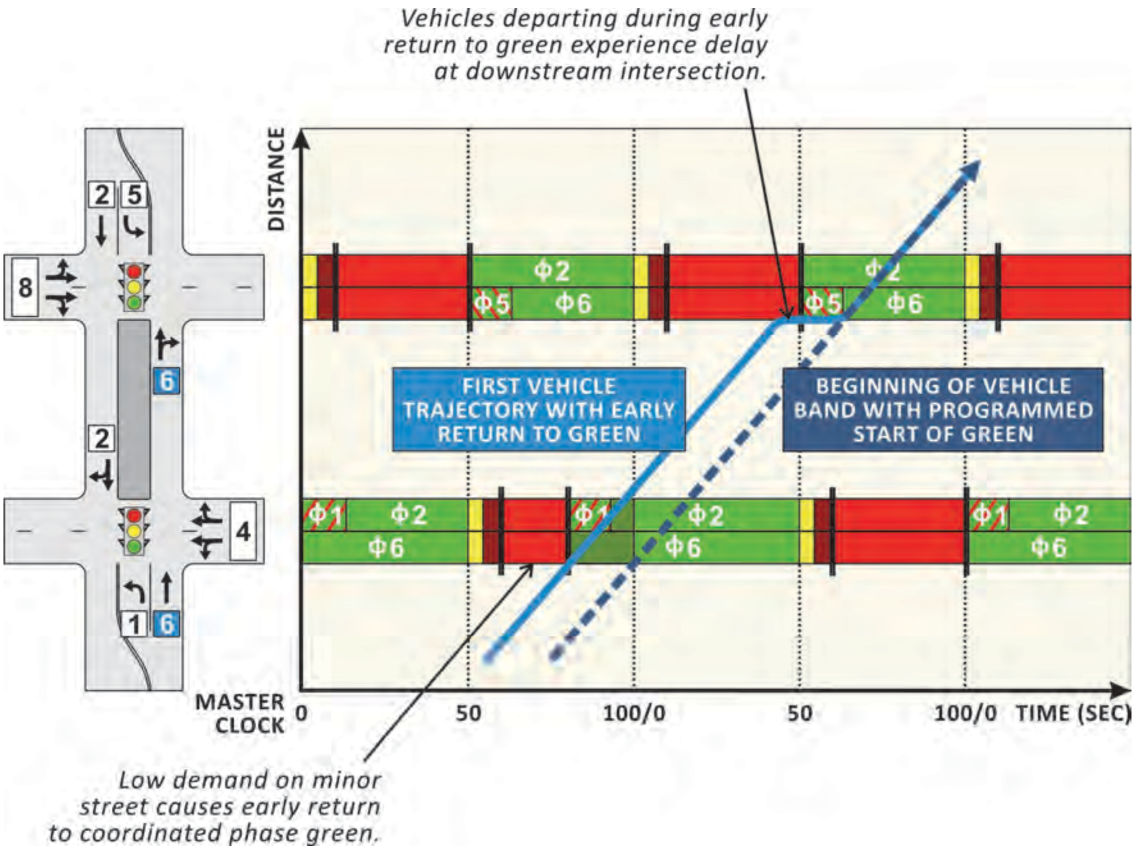
It may be advantageous in some circumstances to adjust the protected left-turn phase sequence for the minor street. In doing this, it may be possible to reduce the delay and queuing for minor street left turns as they enter the major street and arrive at downstream intersections. Such adjustments (which require an optimization tool that shows minor street platoons) may affect system-wide delay, stops, and bandwidth on the corridor.

7.6.2 Early Return to Green

One of the consequences of actuating the uncoordinated phases is the potential for the coordinated phases to begin earlier than expected. This “early return to green” occurs when the sum total of the time required by the uncoordinated phases is less than the sum total of the vehicle splits programmed for the phases. While this may reduce delay at an intersection, it may increase stops at downstream intersections, which is perceived poorly by users. However, if early release does not discharge vehicles or releases queued vehicles ahead of an arriving platoon, it may have positive consequences. It is, therefore, necessary to review several intersections from a system perspective to determine the effects of early return to green.

Exhibit 7-30 illustrates early return to green within a time-space diagram. The figure shows that if the coordinated phases begin early, vehicles may be forced to stop at one or more downstream intersections until they fall within the "band" for that direction of travel. This can result in multiple stops for vehicles and a perception of poor signal timing. Early return to green can be difficult to manage along a corridor, and it can rarely be completely prevented without eliminating most of the benefits of actuation. One technique that can be used is to delay the start of the coordinated phases (i.e., shift the intersection offset to account for the expected early green) if there is a high probability of early return to green. For example, early return to green may be probable if there is a large minor street split (possibly programmed to accommodate an occasional pedestrian call) that is often not fully utilized.

Exhibit 7-30 Example Time-Space Diagram Showing Early Return to Green



7.6.3 Heavy Minor Street Volumes

Heavy minor street volumes can affect the ability to progress through movements along a corridor. These minor street volumes can come from signalized intersections within the coordinated signal system, unsignalized intersections, or driveways between coordinated signals. Interchanges are also a common source of heavy minor street volumes.

In many cases, this additional demand proceeds along the remainder of the corridor and becomes part of the major street through demand at downstream intersections. However, this demand often enters the system outside the band established for through movements traveling end-to-end along the corridor. It may be desirable to adjust downstream intersection timing to allow these heavy minor street movements to proceed with a minimum number of stops, unless it is a minor street with a likely early return to green.

7.6.4 Turn-Bay Interactions

Turn-bay (or turn-pocket) interactions can impact the effective capacity of an intersection. This is experienced when either demand for the turning movement exceeds the available storage space or when vehicle queues block the entrance of a turn bay. Turn-bay overflows can adversely impact progression by disrupting through traffic from proceeding to downstream intersections or altering the arrival profile. This reduces the ability for downstream intersections to efficiently provide green time for the platoon. Beyond being aware of the effects of turn-bay interactions on coordination, practitioners can find techniques to mitigate this type of oversaturation in Chapter 12.

7.6.5 Critical Intersection Control

A challenging aspect of timing an arterial street or a network of streets is the need to provide enough capacity for major intersections without creating excessive delay for minor intersections. Ideally, all of the intersections to be coordinated operate optimally with similar cycle lengths. However, most arterial streets do not have this optimal arrangement due to a mixture of minor intersection signals (e.g., no left-turn phases) with more complex signals (e.g., eight phases), wide ranges in cross-street volumes (e.g., major arterials versus collectors), and variations in left-turn volumes. Several techniques can be used in situations where there is a significant disparity in the ideal cycle length:

- Each intersection is timed using the critical intersection cycle length. This ensures the ability to coordinate all of the intersections in the system. However, use of this technique may result in excessive delay at minor intersections.
- Each intersection is timed to either the critical intersection cycle length or to half that value. This technique is commonly referred to as “double cycling” (i.e., a minor intersection cycles twice as frequently as a major intersection) or “half cycling” (i.e., a minor intersection has half the cycle length of the major intersection). It is also possible in some controllers to have two unequal (asymmetrical) cycles by not providing all phases or constraining splits. These methods can often produce substantially lower delays at the minor intersections where double cycling is employed. However, it may become more difficult to achieve progression in both directions along the major arterial, which could result in more arterial stops than desired.
- The major intersections are operated freely, and the minor intersections are coordinated using a shorter cycle length. Because the major intersections are operating freely, a traditional coordination band is impossible. Therefore, major street vehicles are likely to stop at both the major intersection and at a downstream intersection due to randomness in arrival at and departure from the major intersection. This technique can often result in lower overall system delay at the expense of additional stops along the major street. However, with splits neither constraining green allocation nor driving up cycle length due to pedestrian times, the critical intersection may have fewer phase failures.

7.7 REFERENCES

1. Henry, R. D. *Signal Timing on a Shoestring*. Report FHWA-HOP-07-006, Federal Highway Administration, United States Department of Transportation, 2005.

2. National Transportation Communications for ITS Protocol. *Object Definitions for Actuated Traffic Signal Controller (ASC) Units*. NTCIP 1202, v01.07, 2005.
3. Teply, S. Saturation Flow at Signalized Intersections through a Magnifying Glass. *Proc., 8th International Symposium on Transportation and Traffic Theory*, University of Toronto Press, 1983.
4. Webster, F. V. *Traffic Signal Settings*. Road Research Technical Paper No. 39. H.M. Stationary Office, London, 1958.
5. *Signal Timing Process Final Report*. Contract DTFH61-01-C-00183, Federal Highway Administration, United States Department of Transportation, 2003.
6. Alexiadis, V., K. Jeannotte, and A. Chandra. *Traffic Analysis Toolbox Volume I: Traffic Analysis Tools Primer*. Report FHWA-HRT-04-038, Federal Highway Administration, United States Department of Transportation, 2004.
7. Sunkari, S. R., R. J. Engelbrecht, and K. N. Balke. *Evaluation of Advance Coordination Features in Traffic Signal Controllers*. Report FHWA/TX-05/0-4657-1, Federal Highway Administration, United States Department of Transportation, 2004.
8. Shelby, S. G., D. M. Bullock, and D. Gettman. Transition Methods in Traffic Signal Control. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1978*, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 130–140.
9. Tian, Z. Z., T. Urbanik, K. K. Kacir, M. A. Vandehey, and H. Long. Pedestrian Timing Treatment for Coordinated Signal Systems. *Proc., 2nd International Conference on Transportation and Traffic Studies*, Beijing, China, ASCE, 2000, pp. 533–540.
10. Pohlmann, T., and B. Friedrich. Traffic Signal Transition in Coordinated Meshed Networks. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2192*, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 97–107.

CHAPTER 8
IMPLEMENTATION AND MAINTENANCE

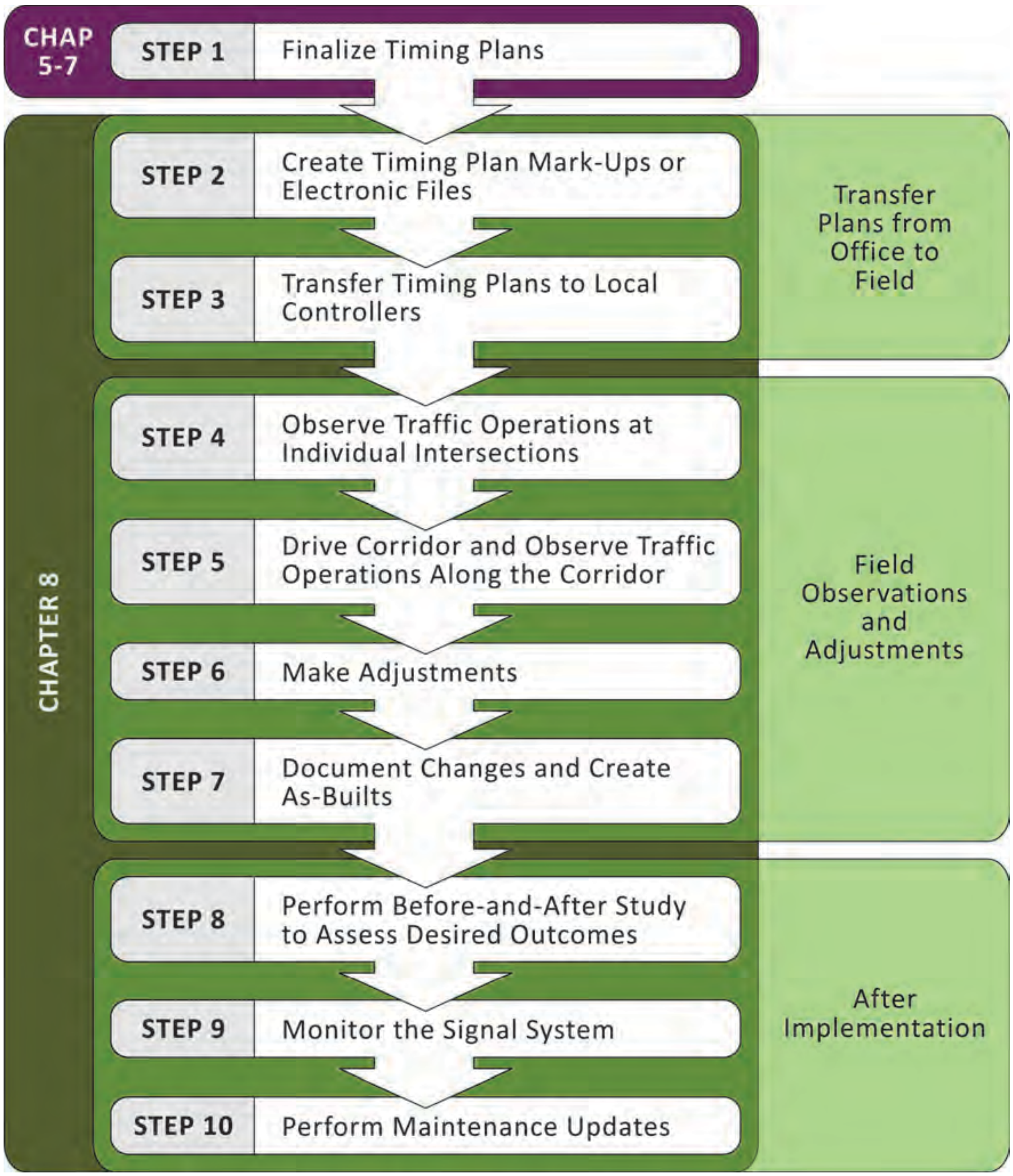
CONTENTS

LIST OF EXHIBITS

CHAPTER 8. IMPLEMENTATION AND MAINTENANCE

This chapter summarizes activities required for effective signal timing implementation and maintenance. Exhibit 8-1 illustrates the process of taking final timing plans (discussed throughout Chapters 5–7) through implementation and to the point where they must be monitored and maintained. Although initial implementation is important for the success of a signal timing project, maintenance ensures that the signal timing will continue to operate at the level expected by the operating agency and general public.

Exhibit 8-1 Signal Timing Implementation and Maintenance Process



Before beginning the process shown in Exhibit 8-1, the practitioner should plan every step with future needs in mind. For example, in order to perform the before-and-after study in Step 8, a practitioner must collect the "before" data prior to implementing the timing plans in Step 3. A procedure should also be developed for saving existing timing plans. Not only will they serve as historical records, but they will also be a reference point if problems are observed after implementation.

8.1 TRANSFER PLANS FROM OFFICE TO FIELD

After timing plans have been finalized in the office, they must be transferred to controllers in the field. Various methods can be used for this transfer depending on the available communication (described in Chapter 4). If a central system has been connected to the field controllers, final plans can be transferred from the office; however, without communication between field controllers or a central system, timing plans must be installed at each location in the field either by hand or electronically.

8.1.1 Finalize Timing Plans

Before implementation can begin, the practitioner should confirm that the elements outlined in Exhibit 8-2 have been defined and approved by the operating agency as part of the signal timing development process.

Exhibit 8-2 Signal Timing Plan Elements

Timing Plan Category	Timing Plan Element	Reference
Objectives	Operating Agency Objectives	3.4 *Operational Objectives and Performance Measures*
General Parameters	Time-of-Day Schedules	6.3 *Time-of-Day Plans*
	Phase Numbering	5.1.1 *Movement and Phase Numbering*
	Phase Sequence	5.1.2 *Ring-and-Barrier Concept* 5.1.3 *Left-Turn Phasing* 7.2.6 *Bandwidth* 7.6.1 *Phase Sequence*
	Overlaps	5.1.4 *Overlaps*
	Detector Assignments	5.1.5 *Detector Assignments*
	Detector Configurations	6.2 *Detector Configurations* 7.5.2 *Actuating the Coordinated Phase*
	Load Switch Assignments	5.1.6 *Load Switch Assignments*
Uncoordinated Parameters	Yellow Change	6.1.1 *Yellow Change*
	Red Clearance	6.1.2 *Red Clearance*
	Minimum Green	6.1.3 *Minimum Green*
	Maximum Green	6.1.4 *Maximum Green*
	Passage Time	6.1.5 *Passage Time (Unit Extension or Gap Time)*
	Minimum Gap	6.1.5 *Passage Time (Unit Extension or Gap Time)*
	Time Before Reduction	6.1.5 *Passage Time (Unit Extension or Gap Time)*
	Time to Reduce	6.1.5 *Passage Time (Unit Extension or Gap Time)*
	Walk Interval	6.1.6 *Pedestrian Intervals* 7.5.1 *Pedestrian Timing and Walk Modes*
	Flashing Don't Walk Interval	6.1.6 *Pedestrian Intervals* 7.5.1 *Pedestrian Timing and Walk Modes*
	Dual Entry	6.1.7 *Dual Entry*
	Recalls	6.1.8 *Recalls and Memory Modes*
	Memory Modes	6.1.8 *Recalls and Memory Modes*

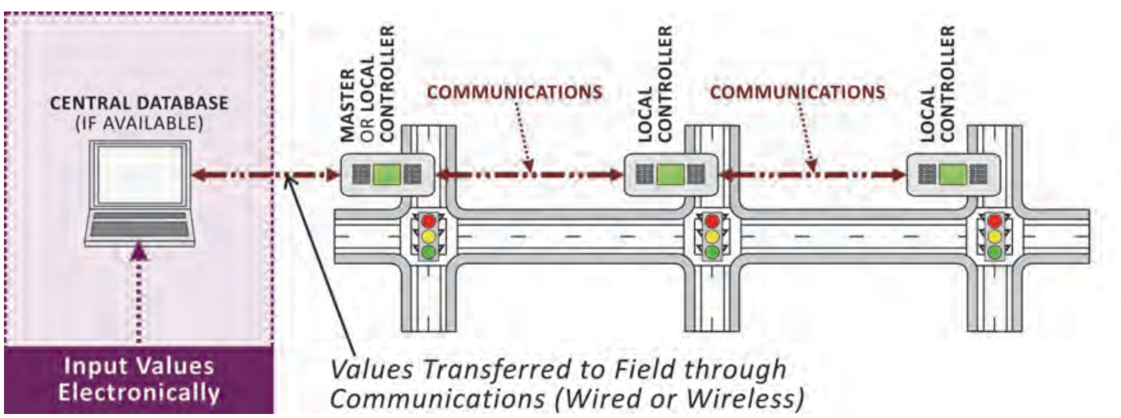
Timing Plan Category	Timing Plan Element	Reference
Coordinated Parameters	Coordinated Phases	*7.3.1 Coordinated Phases* *7.4.1 Coordinated Phases Guidance*
	Cycle Length	*7.3.2 Cycle Length* *7.4.2 Cycle Length Guidance*
	Splits	*7.3.3 Splits* *7.4.3 Splits Guidance*
	Force-Offs	*7.3.4 Force-Offs* *7.4.4 Force-Offs Guidance*
	Permissives	*7.3.5 Permissives* *7.4.5 Permissives Guidance*
	Yield Point	*7.3.6 Yield Point* *7.4.6 Yield Point Guidance*
	Pattern Sync Reference	*7.3.7 Pattern Sync Reference* *7.4.7 Pattern Sync Reference Guidance*
	Offset Reference Point	*7.3.8 Offset Reference Point* *7.4.8 Offset Reference Point Guidance*
	Offsets	*7.3.9 Offsets* *7.4.9 Offsets Guidance*
	Walk Modes	*7.5.1 Pedestrian Timing and Walk Modes*
	Transition Modes	*7.5.3 Transition Logic*

8.1.2 Create Timing Plan Mark-Ups or Electronic Files

As mentioned previously, there are three basic methods for transferring final timing plans from the office to controllers in the field:

1. Use of a central system (as shown in Exhibit 8-3),
2. Electronic transfer in the field (as shown in Exhibit 8-4), and
3. Hand-entering values (as shown in Exhibit 8-5).

Exhibit 8-3 Signal Timing Transfer Using a Central System



If a practitioner plans to use a central system or transfer timing plans electronically in the field, electronic files of the timing plans must be produced. A detailed review of each electronic file should be performed before sending the final files to the controllers. This will minimize the number of data entry errors and allow for adequate preparation prior to field implementation. If yellow and/or red settings are being altered or if phasing is being revised, a formal approval by the operating agency's traffic signal engineer may be required before the changes can be implemented.

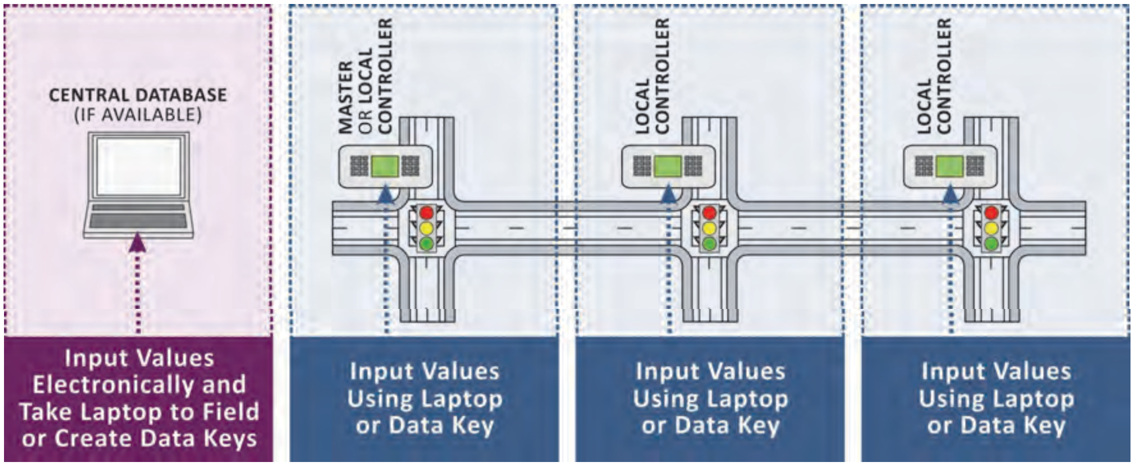


Exhibit 8-4 Signal Timing Transfer in the Field

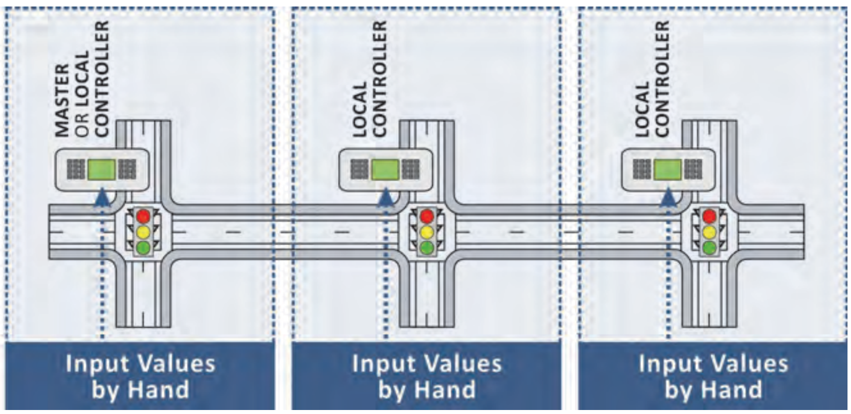


Exhibit 8-5 Signal Timing Transfer Using Hand-Entering

If a practitioner plans to hand-enter values, an easy way to ensure that all of the required signal timing changes are incorporated in the field is to mark-up existing timing sheets by hand, as shown in Exhibit 8-6. Hand-entering values is not the preferred method for signal timing implementation because it increases the potential for errors, but it can be a useful approach when the operating agency does not have the ability to download timing plans electronically. Taking marked-up timing plans to the field will reduce the time required to hand-enter all of the values. Once the draft controller mark-ups are complete, a detailed review of this information should be performed to ensure that the correct timings have been transcribed. Additionally, this review can serve as quality assurance, revealing any parameters that might need to be modified in order for the new timing plans to operate correctly.

Phase Times (next/2/2/2, next/2/2/9/5)

	1	2	3	4	5	6	7	8	
Movement	SBL	NB	WBL	EB	NBL	SB	EBL	WB	
Minimum Green	5	10	5	9	5	10	5	9	0 - 255 sec
Passage	2.0	4.0	2.0	3.0	2.0	4.0	2.0	3.0	0.0 - 25.5 sec
Yellow	4.0	4.0	4.5	4.5	4.0	4.0	4.5	4.5	0.0 - 25.5 sec
Red Clearance	1.0	2.0	1.0	1.0	2.0	1.0	1.0	1.0	0.0 - 25.5 sec or 0 - 255 sec
Max 1	20	40	15	25	20	40	25	25	0 - 255 sec
Max 2	20	40	15	25	20	40	25	25	0 - 255 sec
Walk	0	6	0	10	0	5	0	11	0 - 255 sec
Ped Clear	0	10	0	21	0	15	0	19	0 - 255 sec
Seconds Per Actuation	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0 - 25.5 sec
Time Before Reduction	8	15	6	12	8	15	6	12	0 - 255 sec
Time to Reduce	10	12	10	12	10	12	10	12	0 - 255 sec
Minimum Gap	1.0	2.0	1.0	1.8	1.0	2.0	1.0	1.8	0.0 - 25.5 sec

Exhibit 8-6 Example of Signal Timing Plan Mark-Up

8.1.3 Transfer Timing Plans to Local Controllers

As explained previously, this step depends on the type of communication that is available. Before making any changes, the practitioner should save all existing timing plans in case the new timing plans have unacceptable results. If there is a central system or master controller with communications in place (as illustrated in Exhibit 8-3), the signal timing software will be capable of transferring the signal timing plans directly to each controller. While manual fine-tuning in the field may be required as part of field implementation, manual inputs will be significantly reduced with a central system.

On the other hand, if there is no central system or master controller with communications, the practitioner will need to transfer the timing plans in the field at each individual controller, which can be done electronically (as illustrated in Exhibit 8-4) or manually (as illustrated in Exhibit 8-5). While electronic transfers still take time (because each timing plan must be uploaded individually), this method reduces the potential for human error. Regardless of the method used to transfer the timing plans, the practitioner should verify that the values were successfully uploaded at each controller.

8.2 FIELD OBSERVATIONS AND ADJUSTMENTS

Exhibit 8-7 Field Work

Field implementation (Exhibit 8-7 is practitioner observing real-time operation) is the most critical part of the signal timing process. Both science and finesse are needed to fully realize a good timing plan in the field because of potential site-specific conditions. However, care should be taken not to draw erroneous conclusions from a snapshot of time. Cycle-by-cycle variations will occur, and the practitioner must be patient during observations. A simple and effective means of observation is periodically driving the corridor at different times of day. For example, some staff can be assigned corridors to observe on their way to and from work.

Depending on the history of the corridor and the number of proposed timing plans, field implementation and observations should span a minimum of three days (two days during weekday operations and one day during weekend operations). If feasible, observing traffic operations simultaneously in the field and via traffic management center (TMC) cameras can provide an efficient way to identify problem areas along a corridor (resulting from short split times, excessive queuing, phase failures, and/or hardware issues). The following section provides guidance on field observations and associated signal timing adjustments.

To assist with field implementation, a field notebook should be generated that includes hard copies of the existing and proposed signal timing plans, time-space diagrams, traffic volumes, and a method for noting changes in the field.

8.2.1 Observe Traffic Operations at Individual Intersections

A practitioner should start field observations at individual intersections, ensuring that the timing plans were transferred correctly at each location. There are some signal timing parameters that will be easy to check by timing operations at the intersection, while other parameters will simply need to be verified through observation (as summarized in Exhibit 8-8). In particular, it will be important for the practitioner to determine whether the cycle length and green time distributions are appropriate for the traffic demand.

Exhibit 8-8 Individual Intersection Observations

Timed Parameters	Observed Parameters
☐ Yellow change ☐ Red clearance ☐ Minimum green ☐ Cycle length ☐ Walk interval ☐ Flashing don't walk interval	☐ Phase sequence ☐ Overlaps ☐ Minor street delay ☐ Major street left-turn delay ☐ Vehicle queuing

8.2.2 Drive Corridor and Observe Traffic Operations along the Corridor

The signalized intersections that are part of the field implementation may be operating independently, which would not require a review of the corridor operations. However, if the signal system is coordinated, the practitioner should drive the corridor after observing operations at each individual intersection. A review of vehicle progression can help determine whether changes to the offsets or left-turn phase sequence are necessary.

8.2.3 Make Adjustments

If the practitioner observes opportunities for improvement at individual intersections or along the corridor, he or she can make adjustments in the field. Cycle-by-cycle variations can occur, so the practitioner should ensure that changes are made based on observations of a cross-section of conditions. Exhibit 8-9 provides a summary of some potential adjustments that may improve operations. Because changes can have a domino effect, the practitioner should be aware of how adjustments may influence other movements or intersections. ***Most importantly, the practitioner should keep the operating agency's objectives in mind when making any adjustments in the field.***

Exhibit 8-9 Potential Signal Timing Adjustments Based on Field Observations

Field Observation	Potential Adjustments
Long minor street delay	☐ Redistribute green time between major street phases and minor street phases (e.g., minimum green, maximum green, or splits). ☐ Review cycle length. ☐ Review passage settings for major street phases. ☐ Consider actuating the coordinated phase(s).
Long major street left-turn delay	☐ Redistribute green time to major street left-turn phases (e.g., minimum green, maximum green, or splits). ☐ Review passage settings for major street through phases and minor street phases. ☐ Consider left-turn phase sequence.

Field Observation	Potential Adjustments
Vehicle queuing	☐ Redistribute green time to phases with queuing (e.g., minimum green, maximum green, or splits). ☐ Review cycle length. ☐ Review offsets. ☐ Review passage settings for other phases (not experiencing queuing). ☐ Consider left-turn phase sequence. ☐ Consider phase re-service.
Vehicle platoons arriving on red	☐ Review offsets. ☐ Consider left-turn phase sequence. ☐ Review cycle length. ☐ Review upstream intersections for early return to green and possible offset adjustment.

8.2.4 Document Changes and Create As-Builts

Documenting the signal timing plans that were ultimately implemented in the field is important to the long-term success of a signal system. Changes should be documented in three places:

- ***On Paper.*** Any changes that are made in the field should be documented in both a field notebook that will be taken back to the office and in the cabinet notebook kept at each intersection.
- ***Electronically.*** All electronic files saved to the central computer should be updated at the end of field implementation to reflect the latest timing plans. If a central system and communication exist, the timing plans can be transferred from the field controllers to the central system. In the absence of a central system, a practitioner may want to use a laptop (if available) to transfer timing plans from intersections with changes. However, it may be more efficient to record changes in the field notebook and update the electronic files by hand once back at the office.
- ***In Software Models.*** In order to simplify future maintenance efforts, the practitioner should also update any software used to develop signal timing. If changes in traffic require the operating agency to reconsider operations in the future, a model of final timing values will be ready. Such model networks may also be useful for before-and-after studies, as discussed in the next section.

8.3 AFTER IMPLEMENTATION

After field implementation is complete, there are usually two outstanding items: (1) a before-and-after study (if desired or required) and (2) ongoing monitoring and maintenance. Before-and-after studies are optional, but they can help an operating agency evaluate signal timing objectives and document system performance. Monitoring and maintenance are activities that will continue throughout the life of the signal and are necessary for keeping the signal operating in a manner that is acceptable to the public and the operating agency.

8.3.1 Perform Before-and-After Study to Assess Desired Outcomes

While before-and-after studies are a traditional means of assessment, they take into account a very small sample of operations. Newer techniques of operational monitoring (See Section 8.3.2.1) are beginning to replace more traditional data collection and analysis, such as before-and-after studies, which are often more expensive to perform and more limited in their assessment.

Before-and-after studies commonly compare travel time, delay, and queuing, but an operating agency's objectives should be reviewed when choosing which performance measures to evaluate. While before-and-after studies may incorporate some information from software models, the most valuable information will come from the field. For example, travel time and delay are often evaluated using floating cars that drive the corridor, typically while using a GPS device that records distance and time.

Time-space diagrams (introduced in Chapter 7) are often used to illustrate the before-and-after effects of signal timing plans, specifically those related to travel time, delay, and speed. A time-space diagram can highlight the locations on the corridor where users experience stops and long delays (as shown in Exhibit 8-10). Overall, this example shows that progression along the corridor improved after the new signal timing plans were implemented; the time to travel the entire corridor decreased.

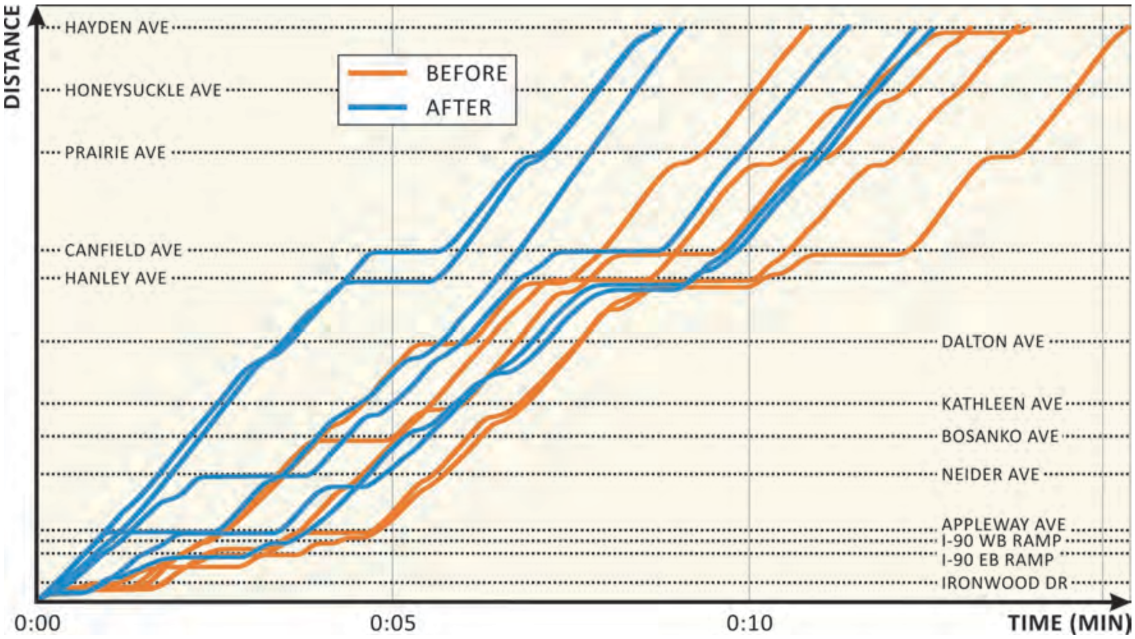


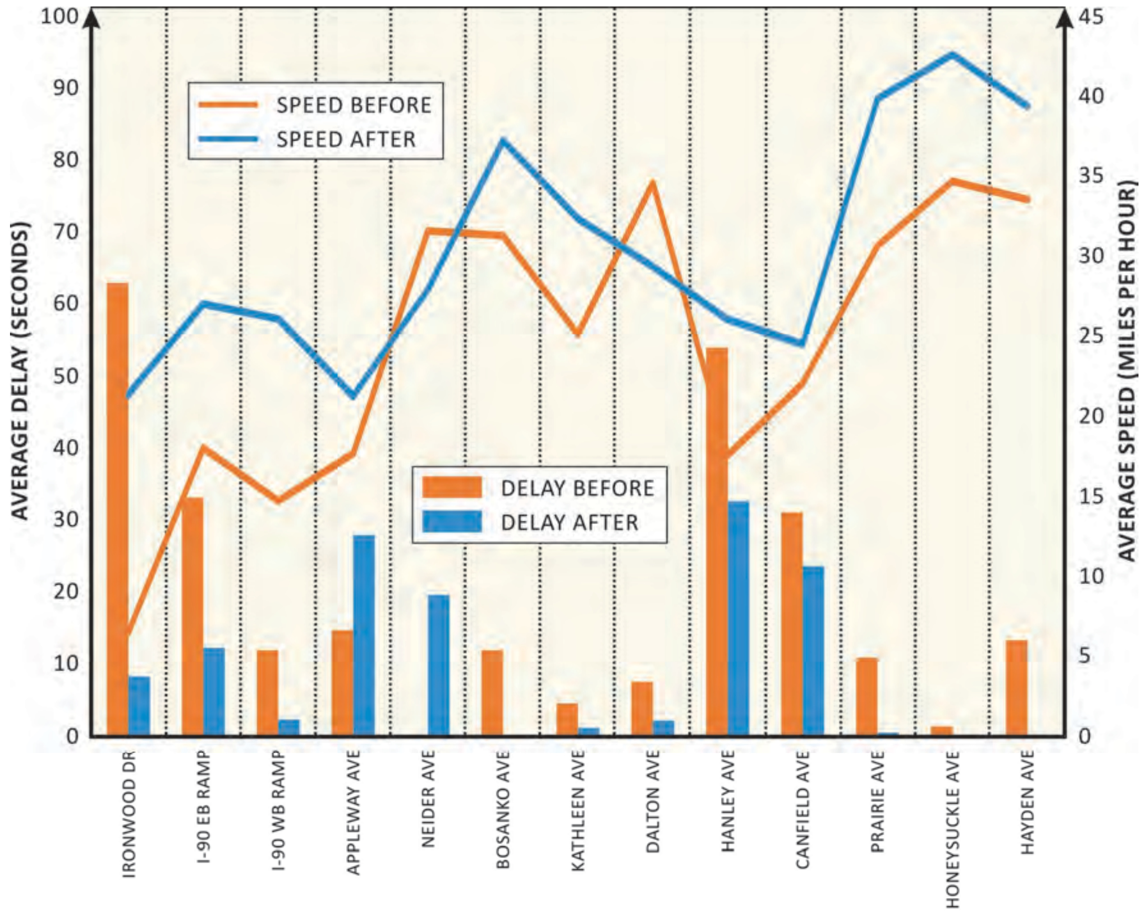
Exhibit 8-10 Example of Before-and-After Comparison: Time-Space Diagram

Graphs of performance measures can also help a practitioner identify changes that resulted from new signal timing plans. Exhibit 8-11 illustrates delay and speeds experienced at specific intersections along a corridor. This example shows that speeds increased after implementation of new signal timing plans, and delay was distributed more evenly across the corridor, with most locations experiencing a decrease.

A before-and-after study report is an opportunity to confirm that the objectives of a retiming effort were met. Some agencies use this step in the process to inform elected leaders about the success of a project and to document future opportunities. It may be posted on the agency website and/or passed along to news agencies for the public to view. The after study is also a benchmark for use in ongoing monitoring of the system and can be used as a basis to determine retiming needs.

When developing a project, it is important to determine the type of before-and-after report that will be needed.

Exhibit 8-11 Example of Before-and-After Comparison: Delay and Speed Graph



Note: Speeds and delays are measured from random arrivals.

8.3.2 Monitor the Signal System

Monitoring a signal system allows an operating agency to proactively make adjustments or respond in a timely and efficient manner to external input (e.g., public service requests). Signal system monitoring should include not only the functionality of signal timing parameters, but also signal system equipment (e.g., detection) that influences intended timing operations. Common monitoring activities, which are described throughout this section, include

- Operational monitoring,
- Equipment monitoring,
- Reviewing changes to agency policies or national standards, and
- Responding to public service requests.

If possible, it is desirable to use data that the signal system automatically collects to assess operational and equipment performance. For example, some signal systems can be programmed to collect data and graphically report various performance measures, as shown in Exhibit 8-12. However, obtaining meaningful real-time performance data from signalized intersections has historically been difficult because detection systems are not designed to provide good count information. Data are often summarized in averages that may not be calibrated, and access to the information is typically limited to personnel managing the signal system.

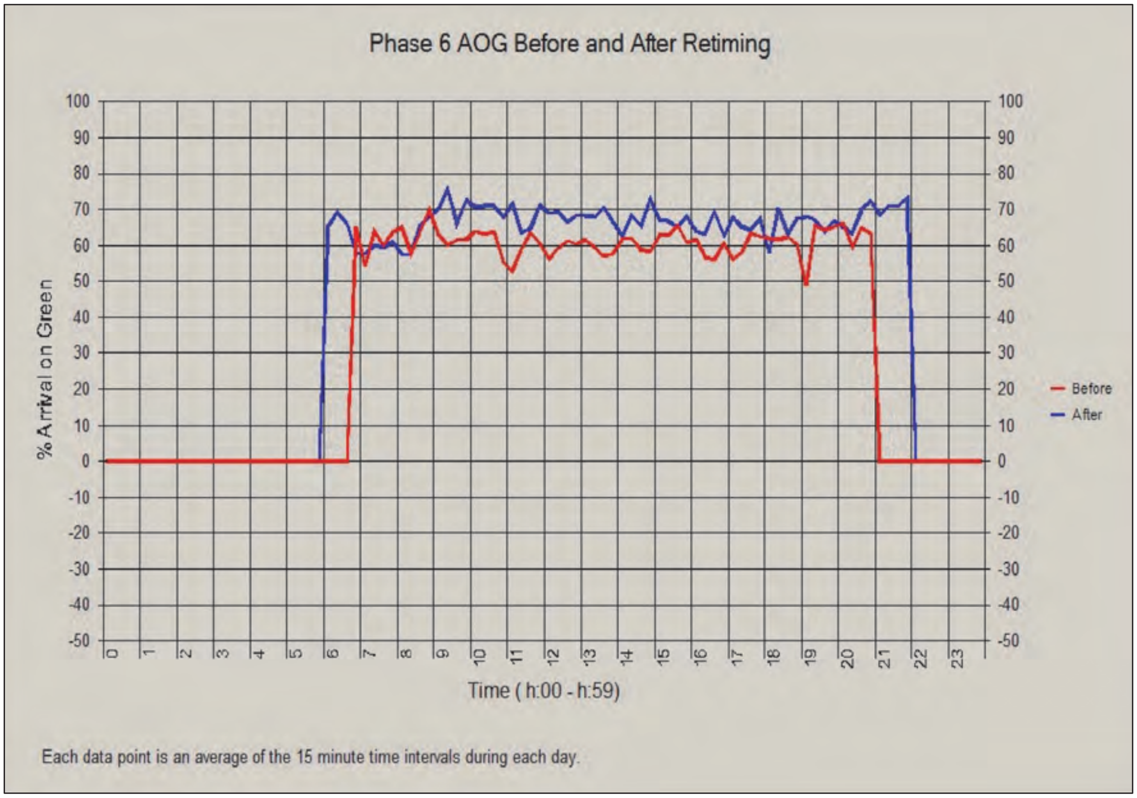


Exhibit 8-12 Example of Before-and-After Comparison: Percent Arrival on Green

8.3.2.1 Operational Monitoring

Operations at a signalized intersection or along a signalized corridor can change for a variety of reasons, including

- Changes in traffic demand (e.g., change in minor street demand, change in turning movement volumes or spillback, change in major street demand, change in vehicle mix, or change in time-of-day volume distribution);
- Changes in roadway geometry (e.g., addition of an approach lane or moving a bus stop from near-side to far-side); and
- Changes in pedestrian traffic due to land use changes (e.g., the opening of a school or residence for the elderly).

Operational monitoring can be conducted through relatively low-tech methods (e.g., signal operators traveling a corridor during their commute) or high-tech methods that use controller or central system functionality to actively monitor performance. Exhibit 8-13 provides examples of low- and high-tech methods for monitoring a signal system.

Exhibit 8-13 Operational Monitoring Activities

Low-Tech Monitoring Activities	High-Tech Monitoring Activities
☐ Scheduled intersection visits ☐ Established commute routes for signal personnel to observe operations during commute ☐ Public service requests ☐ Track growth areas and/or land use changes ☐ Track changes in crash patterns	☐ Mobile applications to monitor intended versus actual operations ☐ Controller or system-based performance measure logs ☐ Permanent corridor travel time data collectors

Several vendors currently support "high-resolution" data logging in the most recent versions of their controllers.

8.3.2.1.1 High-Resolution Data

"High-resolution" event data are an emerging source of data. Rather than storing the average values of data, individual time-stamped traffic events (i.e., when a detector turns on or off or when a phase turns green, yellow, or red) are logged in the controller, or in an external data collector, at a resolution of 0.1 seconds or faster. This type of data can support a variety of performance measures and other applications used to evaluate and improve signal operations (*1*). These performance measures have been successfully used by agencies in Indiana, Utah, Minnesota, and elsewhere.

One application of high-resolution data is a visualization tool called the "Purdue Coordination Diagram" (PCD). The PCD is a useful tool that enables practitioners to quickly evaluate (1) how well a coordinated signal is operating on a particular approach and (2) operations before and after signal timing improvements are made (such as signal retiming or deployment of an advanced control system) (*2*). Exhibit 8-14 is an example PCD illustrating the operations on a coordinated approach over a 24-hour period. Information available in a PCD includes the following:

- The dots represent vehicle arrivals as measured by a setback detector. These are plotted by time of day along the horizontal axis and by time during the cycle along the vertical axis.
- The green and red lines show the start and end of green, respectively, during each cycle. Vehicle arrivals that occur above the green line represent arrivals during a green indication, while those occurring below the green line are arrivals during a red indication.
- The upper red line is also indicative of the cycle length occurring at the intersection. Exhibit 8-14 shows coordinated operations that start at 6:00 a.m. and continue through the end of the day. The cycle length is 120 seconds, with some fluctuation occurring due to actuation of the coordinated phase. Before 6:00 a.m. and after 10:00 p.m., cycle lengths fluctuate considerably during fully-actuated, uncoordinated operations.
- Platoons of vehicles are evident in the clustering of dots that occur throughout the day. Exhibit 8-14 shows that during the morning (6:00 a.m. to 9:00 a.m.), large clusters of vehicle arrivals are evident above the green line, which indicates arrivals on green (or good quality progression). In the afternoon (3:00 p.m. to 7:00 p.m.), this is not the case. The platoons arrive slightly before the start of green, which indicates an opportunity to improve progression (i.e., by adjusting the offsets).
- The same data that are used to produce the PCD can also be used to compute a quantitative performance measure. For example, the percentage of vehicles arriving on green is shown for each coordination plan at the top of Exhibit 8-14.

To demonstrate the utility of the PCD as an evaluation tool, Exhibit 8-15 illustrates operations at the same intersection after offsets have been optimized. During this study, the offsets for all of the timing plans were adjusted, but the afternoon plan (3:00 p.m. to 7:00 p.m.) experienced the most improved quality of progression. In Exhibit 8-15, more vehicle arrivals occur above the green line (i.e., vehicles were arriving on green), illustrating an improvement compared to the "before" case in Exhibit 8-14 (where vehicles were arriving on red). The percentage of arrival on green increased from 53.2 percent to 86.6 percent during the afternoon plan.

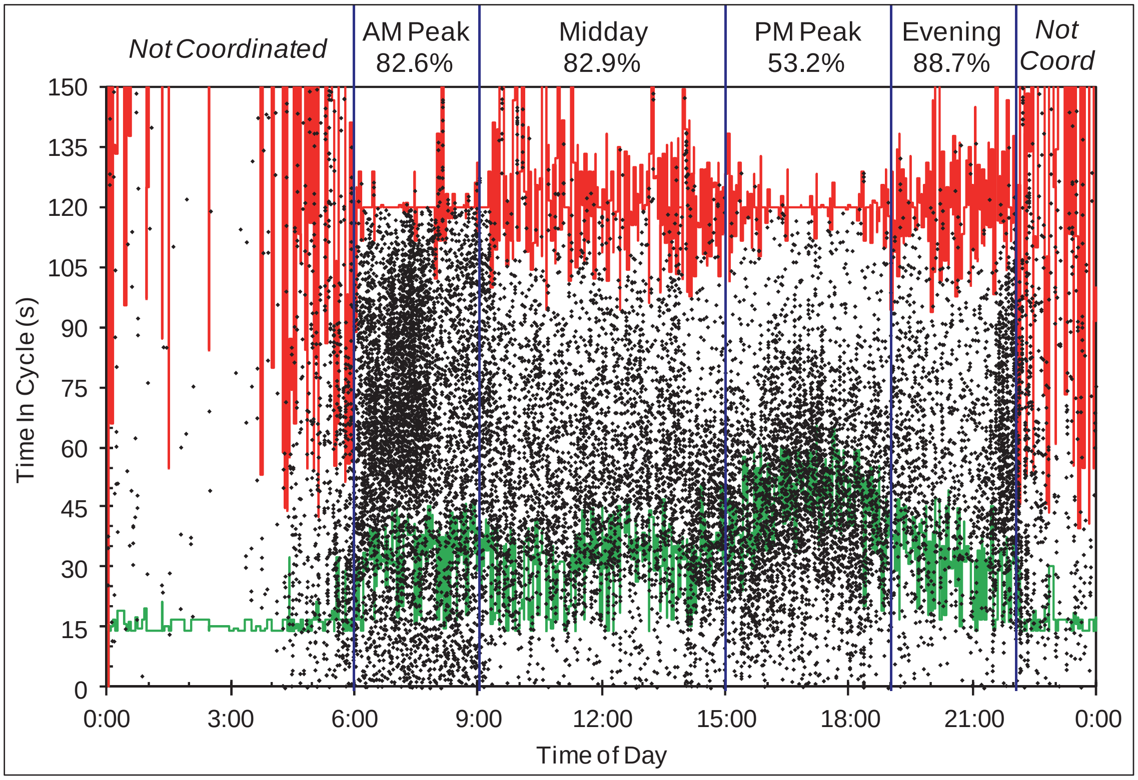


Exhibit 8-14 Example of PCD: Before Offset Adjustment

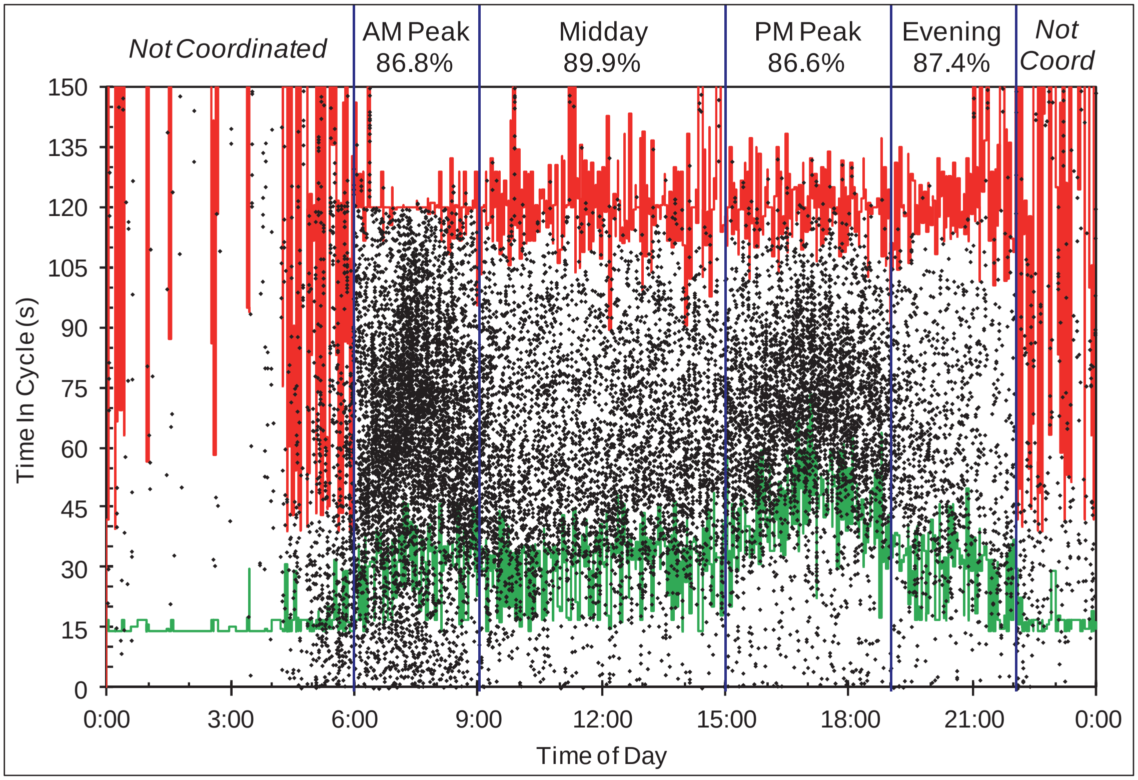


Exhibit 8-15 Example of PCD: After Offset Adjustment

8.3.2.2 Equipment Monitoring

Equipment monitoring is often critical to effective signal timing and signal systems. Equipment failures can range from very noticeable issues (e.g., dark signals or signals in flash) to less noticeable system issues that may cause poor performance but are not as obvious (e.g., clock drift). Exhibit 8-16 lists potential signal-equipment-related issues that may affect signal timing.

Exhibit 8-16 Potential Signal-Equipment-Related Issues

Signal-Equipment-Related Issue	Description	Signal Timing Impact
Detector Fail/ Pedestrian Push Button Fail	Failure indication of a detector (e.g., fail on, fail off, or chatter).	Fail on results in maximum recall, regardless of user demand. Fail off results in user demand not being served, encouraging control violations due to excessive delay.
Count Irregularity	Detected discrepancy in counts on a movement found by comparing recent or real-time counts against an expected profile.	None, unless timing is related to measured volume (e.g., responsive, adaptive, or conditional service).
Controller Status	Online/offline status and/or failure mode indication of a traffic signal controller.	Offline limits remote timing changes, monitoring, log access, etc. Failure modes have varying impacts.
Alarms	Number of controller alarms reported.	Varies.
Controller Parameter Inconsistency	Detected discrepancy in controller parameter settings found by comparing actual and expected controller parameters (likely in a database).	Varies.
Timing Irregularity	Detected discrepancy in timing for a movement found by comparing real-time timing against an expected condition.	Varies.
Clock Synchronization	Check for correct local controller clock setting.	Loss of progression.
Transition Frequency	The number of times and duration that a signal is reported in transition.	Varies, some transition is normal; excessive transition may indicate a problem.
Flash/Conflict Frequency	The number of flash/conflict events, by type if available.	Varies, but typically is very disruptive to traffic.
Power Source	Signal operating on normal or back-up power.	Signal not operating on normal power supply may be in flash.
Communication Quality	The number of successful, failed, and bad poll messages; portion of time a device is connected or disconnected via communication.	Offline limits remote timing changes, monitoring, log access, etc. Failure modes have varying impacts.
Response Time	How long to repair a detected issue.	Varies.
Repair Frequency	Frequency of repair to a consistent issue and/or location.	Varies.

The health of signal equipment should be assessed prior to timing update efforts. For example, if vehicle detection for a left-turn movement is discovered to have failed, the detection should be fixed prior to making signal timing adjustments. In order to streamline signal timing efforts, the operating agency should maintain a database of signal equipment, where technicians can record any signal equipment failures or changes that occur.

8.3.2.3 Reviewing Changes to Agency Policies and National Standards

Agency policies and national standards influence both overarching agency objectives and specific signal timing parameters. A practitioner should monitor changes to policies and standards because they may impact when signal timing updates are made to certain signals and how values are chosen for signal timing parameters.

For example, a city might have a new comprehensive plan that includes a goal to improve bicycle routes around the city. While vehicle delay may have been the primary evaluation tool before, the number of stops for bicyclists could now be an important measure for certain routes, which would require the signal timing to be adjusted.

On a national level, the *Manual on Uniform Traffic Control Devices for Streets and Highways* (MUTCD, *3*) influences specific signal timing parameters. The MUTCD recently changed the walking speed used to calculate pedestrian intervals (from 4 feet per second to 3.5 feet per second), which necessitates an update at all intersections timed under the old standard. It is important for a practitioner to stay apprised of these types of changes so that signal systems remain current.

8.3.2.4 Responding to Public Service Requests

Citizen input is one of the most common reasons for reviewing intersection operations. The public may give input for any number of reasons, including

- A lack of understanding of intersection and controller operations,
- A signal that was in transition between two different timing plans,
- An equipment failure,
- A legitimate observation regarding a shortcoming in the existing timing, or
- An incident near or at the intersection that impacted traffic operations.

Some agencies include a sticker with a logo, phone number, and catch phrase on the outside of the traffic signal controller cabinet to assist the public with obtaining the correct contact information.

Citizens often have a sophisticated understanding of intersection operations resulting from their familiarity with a given roadway. For this reason, as well as for possible reasons of safety, their input should be taken very seriously. Each agency should have a process in place to receive public input and address concerns in a timely, professional manner.

8.3.2.4.1 Collecting Data from the Public

A well-managed operating agency will employ a procedure similar to that below in response to public input (which could arrive by telephone, email, or a letter):

1. Identify the name and contact information of the caller.
2. Identify the location where the problem occurred.
3. Define the time of day when the problem was observed.
4. Determine whether the problem is recurring.

5. Ask for a description of the problem in terms of traffic conditions and traffic signal operations. (For example, if preemption exists at the signal, ask the caller if there were emergency vehicles or rail activity.)
6. Assure the caller that the problem will be investigated within a predefined number of days that has been established by agency policy.
7. Enter all information provided, along with the time and date of the contact, into a database.

Investigation of the problem should be scheduled as part of the agency's maintenance or operations program. If signal timing adjustments are required, the procedures described later in this chapter can be used to address the problem. In all cases, the results of the investigation should (1) be recorded in a database (or written documents) and (2) communicated to the person who made the request using the same media (e.g., telephone, email, or mail) that was used to make the original contact. If no change was made, the reason for maintaining the status quo should be explained.

A website may be used to record input and provide an estimate of response time, as well as to provide a way to maintain the records database. This is a particularly effective technique in regions where multiple jurisdictions are involved with signal operations and maintenance. Establishing a regional website allows problems and public service requests to be directed to the appropriate agency without requiring callers to determine the responsible agency themselves.

Responsive service is the key to good customer relations.

While some agencies utilize sophisticated call-processing software that handles the database functions described here, smaller agencies may find simple spreadsheets to be just as effective for keeping track of public service requests. In either case, it is critical to ensure that all calls are investigated and that a response is provided to the caller in a timely manner. Ideally, a response should be received by the caller within 1 week of the date that the initial contact was made.

8.3.2.4.2 Investigating Service Requests

Diagnose the operational problem before opening the cabinet door, unless there is an immediate safety issue.

After information has been collected from the public about a perceived problem, the first task for the operating agency is to identify whether the reported problem is related to the operation of the signal. Commonly reported issues that are unrelated to signal operations include

- A dark signal due to a power outage,
- A signal with damaged hardware,
- A burned-out bulb, and
- A signal-monitor flash.

Exhibit 8-17 includes examples of common public service requests that are related to signal operations, as well as questions that the operating agency should consider when trying to address these problems. When evaluating a problem in the field, a diagnosis of the potential operational problem should be made before opening the signal cabinet door, unless there is a current safety issue.

Exhibit 8-17 Public Service Requests and Associated Signal Operations Considerations

Signal Operations Category	Example Public Service Requests	Potential Questions to Identify a Solution
Not Getting a Green	☐ "My movement is not getting a green."	☐ Is the intersection part of a coordinated system (that dedicates time to certain phases for progression)? ☐ Have the detectors been damaged? ☐ Is the stopping point well defined (so that vehicles will stop over the detectors)? ☐ Is the detection zone appropriate (e.g., large enough to detect vehicles in a wide approach, sensitive enough to detect bicycles)? ☐ How are the detectors being operated? Is the non-locking setting being used when needed? ☐ Was preemption active?
Short Green	☐ "My movement gets a green, but the green is too short."	☐ If the intersection is part of a coordinated system, is the correct plan running? ☐ Are the splits appropriate and customized for the intersection? ☐ If the particular phase is on recall, are the detectors working properly? ☐ Was preemption active? ☐ If it is a multi-lane approach, is a lane temporarily out of service due to construction or incomplete snow plowing? ☐ Is the approach on a steep grade, where a slippery road could affect performance?
Excessive Delay Before Green	☐ "My movement gets a green, but the delay is excessive."	☐ If the intersection is part of a coordinated system, is the correct timing plan running? ☐ How are offset adjustments being made? ☐ Are detectors located appropriately (so that the first vehicle places a call)? ☐ Was preemption active (potentially for a long time, if an emergency vehicle pulled over with the emitter still running)?
Protected Left-Turn Delay	☐ "When I am making a left turn, why do I have to wait for the arrow when there is no traffic?"	☐ Would protected-permitted left-turn phasing be appropriate? ☐ Should the protected movement be restricted to certain times of day or certain traffic conditions?
Unexpected Stops at a Second Signal After Leaving the First Signal	☐ "I leave one signal, and then I am suddenly stopped at the next signal."	☐ Are the intersections closely spaced and not coordinated? ☐ If the intersections are coordinated, is this condition predicted (such as in the off-peak direction of a one-way progression plan)? ☐ If coordinated, is the first intersection experiencing early release because of light traffic (that could be addressed through a split or offset adjustment)? ☐ How are offset adjustments being made? ☐ If the intersections are part of a coordinated system and on the major street, are the extended detectors at the second intersection working correctly? ☐ If the approach phase at the second intersection is on recall or running semi-actuated, is there a detector malfunction? ☐ Are the intersections part of an adaptive system and in a location with speeds greater than 40 miles per hour (mph)? Depending on the system, there may or may not be setback detection (or even detection directly controlling the length of individual green intervals).

Signal Operations Category	Example Public Service Requests	Potential Questions to Identify a Solution
Signal Serves Approach or Movement with no Demand	□ "I am waiting while an approach with no traffic is served before me."	□ Is there unnecessary locking memory? □ Is the phase on recall? □ Are detection zones misplaced? □ Is there right-turn-on-red with inappropriate detection and delay settings?
Red Light Running	□ "I see a lot of red light running."	□ Are the yellow change and red clearance intervals set appropriately? □ Is the minimum green set appropriately (or will drivers be surprised by a short green)? □ Is the cycle length appropriate? Drivers who are familiar with a signal that has a long cycle length may be encouraged to run the red light instead of waiting through another cycle. □ Are the detectors placed in an appropriate location? □ Do the detectors provide adequate decision zone protection?
Near Misses	□ "I see a lot of near misses."	□ Are the yellow change and red clearance intervals set appropriately? □ Is the minimum green set appropriately (or will drivers be surprised by a short green)? □ Is the cycle length appropriate? Drivers who are familiar with a signal that has a long cycle length may be encouraged to run the red light instead of waiting through another cycle. □ Are the detectors placed in an appropriate location? □ Do the detectors provide adequate decision zone protection? □ If left-turn movements are permitted, should a portion of the movement be protected? □ If left-turn movements are protected-permitted, should the timing and detection for the protected interval be revised?
Short Walk Time	□ "I only get partway into the intersection and the flashing don't walk comes on."	□ Does the caller understand how a pedestrian phase is timed and what the indications mean? □ Is there heavy pedestrian activity or heavy turning vehicle traffic across the crosswalk that would justify a longer walk interval? □ Is the advance walk feature appropriate?
Short Flashing Don't Walk Interval	□ "I cannot finish crossing during the flashing don't walk."	□ Does the caller understand how a pedestrian phase is timed and what the indications mean? □ Does the intersection serve elderly or young pedestrians that would require a longer pedestrian clearance time? □ Does the intersection have heavy turning vehicle traffic across the crosswalk that would justify a longer pedestrian clearance time?
Delay in Getting Walk	□ "I push the button, and it takes a long time to get a walk."	□ Does the caller understand how a pedestrian phase is timed and what the indications mean? □ Are the push buttons operating properly? □ Is the cycle length too long for conditions? □ Can the signal be operated in free mode?
Don't Walk Understanding	□ "I see a circular green, but I don't get a walk (even though I pushed the button). When I start to cross, left-turning traffic nearly hits me."	□ Is split phasing being used (where a pedestrian may be confused about conflicting traffic) without the use of left-arrow displays?

Signal Operations Category	Example Public Service Requests	Potential Questions to Identify a Solution
Emergency Preemption Understanding	☐ "When an emergency vehicle goes through the intersection, the signal operates "funny," and what does the flashing white light mean?"	☐ Would a public information effort be appropriate?
Railroad Preemption Understanding	☐ "When a train goes through the crossing adjacent to the intersection and I am on the other roadway, the delays are excessive."	☐ Is the system operating as designed? ☐ Is the preemption operation efficient?
Railroad Crossing Understanding	☐ "When I am on an approach to a signal with a railroad crossing right at the intersection, I am not sure what I should do."	☐ Is the preemption timing correctly (particularly the track clearance phase timing and the timing of pedestrian intervals)? ☐ Are appropriate signs provided on each approach? ☐ Is detection working appropriately (particularly on the approach behind the tracks)?
Night Flash Operations	☐ "Late at night, the signal creates excessive delay, and there is little to no traffic. Why can't it be in flash?" ☐ "The signal is still in flash in the morning when traffic gets heavy."	☐ What is the agency's night flash policy? ☐ Are the hours of flash set correctly in the controller? ☐ Have the peak periods shifted during the day (requiring flash to be run during different hours)? ☐ Is the time (set in the controller or, alternatively, in the time-switch that controls flash) set correctly?

8.3.3 Perform Maintenance Updates

If deficiencies are identified as part of the monitoring activities explained in the previous section, the following process can be used to make signal timing updates. Because updates are made for a variety of reasons, the following process should only be used as a starting point, with adaptations made on a case-by-case basis.

Before heading to the field to address a signal timing problem, the operating agency should complete the following activities:

- Determine the possible presence of maintenance or construction activity at or near the signal in question.
- Assess the traffic conditions likely to be experienced during the time of day when the problem was observed. A review of available traffic data may be required.
- Review previous retiming reports, if available.

Signal timing updates are different from the systematic retiming efforts discussed in Chapters 5–7, but the field work will utilize a procedure similar to that used for implementation of a new timing plan (described previously in this chapter).

- If the intersection has remote monitoring capabilities, the problem should be investigated with that tool first. In particular, performance monitoring logs may be helpful.

If the operating agency is not able to determine a solution based on office activities, a field visit may be warranted, and the following procedure can be employed:

- Schedule a field visit for the time of day and type of day (e.g., weekend or weekday) during which the problem was identified.
- Assemble the timing and configuration information for the intersection being visited. Timing information should include controller settings and, if available, traffic count data. If the intersection is part of a system, the information should also include the master clock, offsets, and time-of-day schedules.
- If the intersection is included in a system, coordinate with system operators to ensure that operations personnel will be available to support the field activities.
- When arriving at the intersection, observe the physical condition of the street hardware (including the poles, mast arms or span wires, signal heads, detection, pedestrian displays, and cabinet). Make a preliminary assessment of the existence of a problem and its likely cause *before* opening the cabinet (unless an obvious safety problem needs to be immediately addressed).
- Open the cabinet and check the log book, and then perform a physical inspection of the cabinet interior (including cabling, physical condition and operation of cabinet components, air filter, and fan).
- Check operability of all cabinet components either through observation or suitable maintenance diagnostics.
- Review controller timing by comparing settings with timing documentation.
- Qualitatively compare traffic conditions at the intersection with the traffic count data (data must be for same time period). Determine whether major changes in demand have occurred since the traffic counts were taken to support the development of the timing plans currently in use. If major changes have occurred, determine whether they are temporary (e.g., due to nearby construction) or permanent. If they are temporary, it may still be desirable to update the intersection timing. However, include a note in the maintenance log (and any other cabinet documentation) that a second signal timing update may be required when construction is complete.
- In all cases, after timing and/or scheduling changes have been made, the impact of the changes should be evaluated through observation of the intersection operations.
- The final step of this process is to log the actions taken. This is essential for responding to the individual who initiated the maintenance, as well as records that must be maintained by the operating agency for a variety of engineering, operational, and legal reasons.

8.3.3.1 Common Timing Updates for Under-Saturated Conditions

It is likely that the most frequent requests for intersection maintenance will occur during normal flow conditions. These are the conditions that impact a large number of roadway users, who all expect high signal timing quality. Users will be annoyed at instances of wasted green time when they are waiting at a red signal indication with no vehicles on other phases or when they have to stop at successive signals due to poor offsets.

When retiming an intersection for under-saturated (free-flow) conditions, the following steps are recommended:

- Perform a qualitative evaluation of the intersection performance to determine whether any obvious improvements are possible.
- Adjust the splits to reflect demand on competing approaches.
- Adjust the offsets to reflect platoon arrival times.
- Adjust the start and end times of timing plans (in the time-of-day plan).
- Review the cycle length to determine the need for a new timing study.

8.3.3.2 Common Timing Updates for Oversaturated Conditions

It is important to recognize the need for different traffic signal timing strategies for networks that experience oversaturated traffic conditions (i.e., demand exceeds intersection capacity). Strategies begin to change from mobility and progression to queue management. When initially assessing signal timing at a congested intersection, the splits (first) and cycle length (second) should be reviewed and adjusted if possible to reduce the congestion. (Note that congestion may exist for some users because of higher-priority operational objectives for other users.) If typical under-saturated adjustments are not able to produce reasonable results, then oversaturated techniques (discussed in Chapter 12) may be necessary. However, the presence of congestion should not ***immediately*** cause a practitioner to assume that there is excessive demand. In many cases, congestion is the result of poor signal timing parameters, especially excessively long cycle lengths or bad offsets that do not promote progression.

Congestion can be recognized by the presence of queues (at signalized intersections) that are not completely discharged during the green period. Chapter 12 has additional information on symptoms of oversaturation.

8.4 STAFFING NEEDS

An agency may need a variety of staff positions and roles to adequately operate and maintain its traffic signal system. The size and breadth of the staff dedicated to traffic signal operations and maintenance should depend primarily on the size and complexity of the system.

8.4.1 Staff Positions

Larger organizations have specialized staff to support signal systems. Depending on the size of the signal system, some of these positions may be combined. The roles of each position described below are based on information from agencies and relevant Institute of Transportation Engineers (ITE) and Federal Highway Administration (FHWA) literature:

- ***Traffic Signal Engineer:*** This staff person is responsible for the day-to-day operations of the signal system. Tasks include responding to public comments, approving new signal turn-on’s, assisting in the TMC, evaluating signal timing on existing arterials, managing signal operations staff, and coordinating with the signal design and maintenance supervisors.

The important point is that a complete set of skills is needed to operate and maintain a traffic signal system, but those skills can reside in different combinations of staff positions.

- ***Traffic Signal Technician/Analyst:*** This staff person assists the traffic signal engineer with his or her day-to-day operations. Focus areas include signal timing, new signals, and the TMC.
- ***Intelligent Transportation Systems (ITS) Engineer:*** This staff person is responsible for the implementation of ITS projects. Tasks include responding to public comments, evaluating new products, assisting in the TMC, managing ITS contractors and vendors, and coordinating with the signal design and maintenance supervisors.
- ***Traffic Signal Maintenance Technician:*** This staff person is generally responsible for troubleshooting and maintenance of the physical traffic signal equipment.
- ***Electronic/Communications Specialist:*** This staff person is responsible for the complex electronic equipment at the heart of the signal system. Some tasks include closed-circuit television (CCTV) system repair (field and central system); fiber-optic-cable system testing, repair, and termination; telecommunications system maintenance and repair; TMC system maintenance and repair; traffic signal controller electronics testing, repair, and inventory; and other ITS devices repair. The skills required include knowledge of Ethernet communications, databases, troubleshooting software, and an understanding of the software that monitors, maintains, and runs the system. Information Technology (IT) skills are a vital part of a successful system.
- ***TMC Operator:*** This staff person is responsible for observing traffic conditions, responding to incidents that occur in the field, and providing support to homeland security efforts. This role is critical to the rapid response and resolution of signal system situations.
- ***Public Relations Coordinator:*** This staff person is responsible for fielding phone calls from the public, coordinating with the traffic signal engineer and technician/analyst on responses, and marketing the TMC, incident management plan, and traffic signal operations to the public. Depending on the size of the agency, this position could be a full-time position or these tasks might be passed on to the traffic signal engineer and technician/analyst.

Proximity to traffic signals is an important consideration when determining staffing needs, particularly for technicians who are tasked with responding to public service requests or maintenance problems. A city with a large downtown may have 100 signals within 1 square mile, as opposed to a rural district with 100 signals over 1000 square miles. Obviously, in these cases, the staffing needs are likely different and may require different skillsets and staffing levels.

Additional information regarding staffing needs has been summarized in several publications, including

- **ITE *Traffic Engineering Handbook* (*4*) and ITE *Traffic Control System Operations: Installation, Management, and Maintenance* (*5*).** These texts suggest labor requirements of 20 to 25 hours per intersection for traffic signal retiming. They estimate that one traffic engineer is needed to properly operate and maintain every 75 to 100 signals and that one technician is needed to properly operate and maintain every 40 to 50 signals. While good rules of thumb, the current transportation environment requires much more detailed estimates.

- **FHWA *Traffic Signal Operations and Maintenance Staffing Guidelines* (*6*).** This publication provides staffing and resource guidance for agencies in order to help them effectively operate and maintain their traffic signal systems. Surveys performed by FHWA as background for these guidelines indicate that agencies are able to achieve a high level of performance under a wide variety of signal system conditions. Because of the wide variety of successful staffing and resource combinations, performance-based criteria were developed to help agencies define realistic and concise objectives and performance measures, which can then be used to estimate staffing and resource needs.

8.4.2 Staff Training

A valuable component of traffic signal maintenance is ensuring that the staff managing and maintaining the traffic signals has been trained to operate the system. Training can include activities such as peer exchanges (within or between agencies), technical sessions with outside experts, or attending conferences, educational seminars, or universities.

Critical training elements highlighted in the FHWA *Guidelines for Transportation Management Systems: Maintenance Concept and Plans* (*7*) include

The equipment and software utilized by many agencies for their traffic signal systems are only as good as the availability of skilled and trained staff.

- ***Training by Vendors.*** Procurement contracts should include a requirement for on-site training of agency staff in maintenance and operation of the equipment, preferably conducted by the vendor.
- ***Training by Contractors.*** Procurement contracts should also include a requirement for on-site training of agency staff in the maintenance and operation of the assembled systems, including software, hardware, and devices.
- ***Training Library.*** The operating agency should maintain a library of system documentation and, if available, a video library of training materials.
- ***Staff Retention.*** This can be difficult in a high-tech environment, but there are ways to improve retention, such as supporting additional training, allowing travel to technical conferences and workshops, and providing other benefits for agency staff.

In addition to internal staff training, the operating agency might also consider hosting short work sessions or seminars that are open to the public and other agency officials. External training gives roadway users the opportunity to gain insight into signal timing and traffic operations, which only enhances the external input received for monitoring and maintenance activities.

8.5 REFERENCES

1. Day, C. M., D. M. Bullock, H. Li, S. M. Remias, A. M. Hainen, R. S. Freije, A. L. Stevens, J. R. Sturdevant, and T. M. Brennan. *Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach.* Purdue University, West Lafayette, Indiana, 2014.

2. Day, C. M., R. Haseman, H. Premachandra, T. M. Brennan, Jr., J. S. Wasson, J. R. Sturdevant, and D. M. Bullock. Evaluation of Arterial Signal Coordination: Methodologies for Visualizing High-Resolution Event Data and Measuring Travel Time. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2192*, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 37–49.
3. *Manual on Uniform Traffic Control Devices for Streets and Highways,* 2009 Edition. United States Department of Transportation, Federal Highway Administration, Washington, D.C., 2009.
4. *Traffic Engineering Handbook*, 6th Edition. Institute of Transportation Engineers, Washington, D.C., 2009.
5. Giblin, J. M., and W. H. Kraft. *Traffic Control System Operations: Installation, Management, and Maintenance*. Institute of Transportation Engineers, Washington, D.C., 2000.
6. Gordon, R., and C. Braud. *Traffic Signal Operations and Maintenance Staffing Guidelines*. Report FHWA-HOP-09-006, Federal Highway Administration, United States Department of Transportation, 2009.
7. Vick, C., and R. Sumner. *Guidelines for Transportation Management Systems: Maintenance Concept and Plans*. Report FHWA-OP-04-011, Federal Highway Administration, United States Department of Transportation, 2002.

CHAPTER 9
ADVANCED SIGNAL SYSTEMS

CONTENTS

LIST OF EXHIBITS

CHAPTER 9. ADVANCED SIGNAL SYSTEMS

While earlier chapters largely address “standard” signal timing features, which are fairly similar across all modern traffic signal controllers, Chapter 9 describes advanced signal systems. These systems are able to make signal timing adjustments based on detection information, thus modifying operations during varying traffic flow conditions. Applications can range from the use of advanced coordination features to the deployment of adaptive signal control technology (ASCT).

While there are many variations, most advanced systems require a considerable investment from local agencies. As such, this chapter begins by describing a process known as “Systems Engineering” that can be used to determine whether an advanced signal system is most appropriate for a given location based on local needs and requirements. The following sections focus on several types of advanced signal systems that can be applied, including advanced coordination features, traffic responsive systems, and ASCT systems.

9.1 SYSTEMS ENGINEERING

The main objectives of the Systems Engineering process are identifying and defining an agency’s needs in concrete terms so that a signal technology can be selected that meets those needs. It is recommended that an agency completely assess the capabilities of its existing system before considering alternatives. This ensures that the capabilities of the current system are maxed out before an upgrade is considered. Most advanced signal systems are a considerable investment, so this process is intended to minimize risk for an agency throughout the design and implementation of a new system. Systems Engineering ultimately helps an agency develop five reports that document the decision-making process:

- Concept of Operations (ConOps)
- System Requirements
- Design and Implementation
- Verification Plan
- Validation Plan

While Systems Engineering is often used to assess advanced signal systems, the process can be applied in many situations (e.g., evaluating preferential treatment options).

Systems Engineering was initially developed to design and implement large-scale projects, so the level of effort required to complete the full method is significant. The practitioner should tailor the level of application to the size and/or complexity of the project. The intent of this section is to provide a high-level overview of the Systems Engineering process, highlighting some of the key elements. For additional information and detailed guidance, please see *Systems Engineering for Intelligent Transportation Systems* (*1*) and *Model Systems Engineering Documents for Adaptive Signal Control Technology Systems* (*2*).

9.1.1 Concept of Operations

The Concept of Operations (ConOps) is a ***non-technical*** document that defines a system’s needs (i.e., what problem needs to be addressed) and, more importantly, describes ***how*** the system will be ***used***. The project needs in this document will eventually be connected to technical requirements; all elements of a project are ultimately defined by and referenced back to the ConOps. It is the foundation upon which the rest of the Systems Engineering process builds.

The ConOps is written to, and is the responsibility of, the stakeholders of the project. Typically, stakeholders include anyone impacted by the project and may encompass owners, operators, maintainers, and users. Identifying appropriate stakeholders, clearly articulating project needs, and obtaining consensus on the project needs are important activities in this stage of Systems Engineering.

9.1.2 System Requirements

The System Requirements document is a ***technical*** report written for technical staff, vendors, users, and system operators. The elements within this document link directly back to specific needs in the ConOps and describe ***what*** needs to be ***achieved*** with the project. The requirements do ***not***, however, describe how the system will be built. Exhibit 9-1 summarizes various requirement categories that should be considered, sourced from *Model Systems Engineering Documents for Adaptive Signal Control Technology Systems* (*2*).

Exhibit 9-1 System Requirements Categories

Requirement Category	Description
Functional Requirements	What the system is to do
Performance Requirements	How well it is to perform
Non-Functional Requirements	Under what conditions it will perform
Enabling Requirements	What other actions must be taken in order for the system to become fully operational
Constraints	Limitations imposed on the design by agency policies and practices, such as type of software, type of equipment, and external standards
Interface Requirements	Definitions of the interfaces between subsystems or with external systems
Data Requirements	Definitions of data flow between subsystems or with external systems

Composition of the System Requirements document is the responsibility of the same stakeholders responsible for the ConOps. During development of the System Requirements document, stakeholders must select requirements as well as analyze, document, validate, and manage the requirement statements (within the System Requirements document and how they are referenced to the ConOps).

9.1.3 Design and Implementation

Up to this point in the Systems Engineering process, the activities have focused on defining the problem to be solved but not on the solution. This is the purpose of the design activity. This portion of the process may include conducting alternatives analysis and product evaluations, documenting interfaces and standards, and developing detailed designs and specifications. Implementation of the project involves activities associated with installation, transition, and delivery of the project components.

9.1.4 Verification Plan

Once the project is designed and implemented, a Verification Plan is used to confirm that all requirements (outlined in the System Requirements document) have been met. The Verification Plan is typically a ***technical*** report written for technical staff, vendors, users, and system operators. It describes ***how*** the system will be ***tested*** and documents the activities and results of the verification process. The Verification Plan does ***not*** provide detailed descriptions of data collection or analysis techniques, as that is part of the validation process (described next). The Verification Plan is typically the responsibility of the system developer, vendor, or supplier, and is overseen by the agency.

9.1.5 Validation Plan

The Validation Plan is written after the verification process is complete and describes ***how*** the performance of the system will be ***measured***. The plan defines the measures of effectiveness (MOEs), the data requirements and collection procedures, and the type of analysis required to validate the system against all needs identified in the ConOps. It also documents validation activities, results, and any necessary corrective actions. The Validation Plan is typically developed by the agency and is written for the stakeholders of the project.

9.2 ADVANCED COORDINATION FEATURES

In most cases, standard coordinated signal timing plans sufficiently meet the needs of a corridor or network. However, there are times when standard features are no longer able to respond efficiently to traffic demand or fluctuating patterns. When these situations occur, a common response is to invest in new signal technology, such as ASCT. While an ASCT system may benefit signal operations, the practitioner should first investigate available advanced coordination features and detection upgrades that may address local problems without the substantial investment required by an ASCT system.

The nomenclature and functionality of advanced features differs from vendor to vendor.

This section highlights advanced coordination features typical of most signal systems, but it is not intended to be comprehensive. It is an introduction to select features, and highlights situations where these coordination features are both available and applicable. Some traffic signal controllers will have advanced coordination features not discussed in this section.

9.2.1 Actuating the Coordinated Phase

Actuating the coordinated phase (or fully-actuated coordination) is a signal timing treatment that actuates a portion of the coordinated phase split. Using this feature, the coordinated phases are permitted to gap out during the later portion of their split time in the event of low demand. This allows additional time to be given to the minor street and left-turn phases. Additional information about this particular advanced coordination feature is provided in Chapter 7, as its application is becoming more typical.

9.2.2 Dynamic Phase Length

Dynamic phase length (also known as extended split or adaptive split) allows controller firmware to alter the length of a phase through dynamic movement of the force-off point. This feature allows effective use of unused green time within a cycle and can reduce the queue lengths experienced at an intersection. Note that this feature does not alter cycle length or shorten the coordinated phase(s), so it will not disrupt coordinated operations.

In order for an intersection to benefit from the dynamic phase length feature, a system must be capable of estimating vehicle demand. Vehicle demand can be measured and forecast using queue information from vehicle detectors. It is important to ensure that detection is functional and passage time properly set (particularly at setback detectors located upstream of queued vehicles) when using dynamic phase length.

9.2.3 Phase Re-Service

The phase re-service feature (also known as repeated phase service or conditional service) allows a phase to be served more than once during a cycle. Most benefits of phase re-service are realized during low to medium traffic volume periods, such as the shoulder of peak periods or midday periods. Phase re-service can also be a useful treatment at minor intersections in a coordinated system where the system cycle length is longer than what is needed to serve local traffic demand (and half-cycling is not an option). Engineering judgment and field observation are essential when determining whether phase re-service will benefit a system.

9.3 TRAFFIC RESPONSIVE PLAN SELECTION SYSTEMS

Traffic responsive plan selection systems are able to change coordination plans to fit traffic conditions. Based on detector data returned from the field, traffic responsive algorithms select a preconfigured signal timing plan from a library of plans. The current timing plan must usually have been running for a minimum amount of time before a new plan can be implemented, and the new timing plan must typically be a certain percentage improvement over the currently running plan. If those conditions are met, the signals in a group are directed to begin using the new plan at the same time, in order to ensure that coordination is reestablished quickly.

Most traffic signal system vendors offer a software module (whether standard or optional) that will operate a traffic responsive plan selection system.

Most, if not all, traffic signal systems allow traffic responsive plan selection to be implemented on a time-of-day basis. In this case, the schedule includes a baseline plan that is initiated during a corresponding time-of-day period. If detection data indicate that the scheduled plan is responding inefficiently to the observed traffic patterns, the system switches the scheduled plan to a more suitable alternative. Some systems allow the traffic responsive mode to be directed manually or to be scheduled as a special event override of the normal time-of-day schedule.

Signal timing plans for traffic responsive systems are typically configured to cover a wide range of field scenarios. It is not necessary to have a fully comprehensive set of timing plans for traffic responsive systems to operate. However, all timing plans and pattern data must be configured in the controller prior to field operations. Traffic responsive plan selection systems do not calculate new plans (or patterns).

9.3.1 Traffic Responsive Algorithms

There are two primary categories of traffic responsive algorithms that are used to select timing plans: (1) algorithms that use targets and (2) algorithms that use thresholds. Target-based methods are typically available for NEMA and 170/2070 firmware, while threshold-based methods are typically only available for NEMA controller firmware. Modern NEMA controllers now support both plan- and pattern-based operations.

9.3.1.1 Target-Based Systems

Target-based systems were developed by the Federal Highway Administration (FHWA) in the early 1970s during the development of the Urban Traffic Control System (UTCS).

Target-based systems transition to new signal timing plans based on volumes detected in the field. These systems require a practitioner to program the timing plans that will be applied during certain traffic conditions (defined by target volumes at each detector). The systems then calculate the difference between real-time traffic volumes and the target volumes (or "signatures") preselected for each detector. Based on the relationship between the real-time conditions and predetermined signatures, the central system or master controller applies the timing plan that is most appropriate for the field conditions.

Typically, target values for each detector are expressed as equivalent hourly volumes. Because detector data yield a combination of volume and occupancy data, Equation 9-1 must be applied to compare real-time volumes to predetermined target volumes. The timing plan (and, more specifically, the associated signature) that produces the lowest value in Equation 9-1 is compared to the value produced by the currently running plan. If the yielded value is lower than a percentage of the current value, the new timing plan is implemented.

Equation 9-1

$$err_j = \sqrt{\sum_{i=1}^{N} w_i\big((V_i + KO_i) - (V_j + KO_j)\big)^2}$$

where

- err = difference between field volume and target volume (vehicles per lane per hour),
- N = number of detectors,
- i = field detector,
- j = target detector,
- w = weighting factor for detector,
- V = volume (vehicles per lane per hour),
- K = weighting factor for detector occupancy (vehicles per lane per hour per percent occupancy, typically a global parameter for all detectors), and
- O = detector occupancy (percent).

Detector occupancy is the percentage of the analysis period that a vehicle was stopped on the detector.

In free-flow conditions, the occupancy measured by a system detector is often low and does not contribute to the calculation. However, when traffic conditions are congested and queues form on the detector, detector volume measurements will be constant and much lower than actual traffic demand. Without the use of the KO factor in this scenario, the algorithm logic inaccurately assumes that volumes are low. The KO factor provides a way to scale the occupancy measurement to an equivalent volume.

A rule of thumb is simply to use a K value that scales the occupancy (reported as between 0 and 100 percent) into an equivalent hourly volume. For example, a K value of 17 would scale an occupancy value of 100 percent to 1700 vehicles per lane per hour. A K value of this magnitude might be appropriate for detectors at the exit side of an intersection that infrequently experiences substantial occupancy. For detectors closer to the stop bar that experience more frequent queuing, lower K values (i.e., approximately between 5 and 7) should be used, since occupancy will frequently average non-zero values.

In many cases, detector data are weighted to scale the importance of the information from one detector as compared to that of another. In practice, weights are often used to amplify small changes in volume when a detector does not experience variability as high as other detectors in the system. That being said, there is limited guidance on the configuration of detector weights.

9.3.1.2 Threshold-Based Systems

Threshold-based systems calculate which timing plan (or signal timing parameters) to select based on field detector data crossing certain threshold values. Most threshold-based systems use computation channels (CCs), which are parameters associated with groups of field detectors (e.g., a set of inbound and outbound detectors). In general, CC parameters are the sum of weighted detector data and can be calculated using the general form of Equation 9-2. However, each vendor will use different types and numbers of CC parameters, as well as different weights and smoothing factors, and will apply different combinations of maximums and averages to compare to the timing plans (or parameters). The practitioner should refer to the specific manufacturer for details about CC parameter calculations.

Equation 9-2

$$CC_i = \sum_{i=1}^{N} w_i(V_i + KO_i)$$

where

- CC = computation channel (vehicles per lane per hour),
- N = number of detectors,
- i = field detector,
- w = weighting factor for detector,
- V = volume (vehicles per lane per hour),
- K = weighting factor for detector occupancy (vehicles per lane per hour per percent occupancy, typically a global parameter for all detectors), and
- O = detector occupancy (percent).

The configuration of weights can be particularly important in order to match the field data with appropriate control parameters. A detailed methodology for the configuration of threshold-based systems can be found in *Methodology for Determination of Optimal Traffic Responsive Plan Selection Control Parameters* (*3*). This methodology uses statistical analyses (i.e., discriminant analysis and principal component analysis) of the detector data to identify reliable weights for the linear combination of volume and occupancy data.

After CC values have been calculated for various groups of detectors, they are compared to the threshold values set by the practitioner. Different CC values will trigger different signal timing plans (or parameters including cycle length, splits, and offsets). Rather than using a single target value, threshold-based systems allow the practitioner to assign CC values that the algorithm will use to both enter into and exit out of new timing plans. Exhibit 9-2 illustrates example CC thresholds for the selection of an offset value. In this example, if the CC value is currently 52, then Offset 3 would be the current setting. If the CC value falls below 49, then Offset 2 would be selected, and if the CC value rises above 68, then Offset 4 would be selected.

Exhibit 9-2 Example Traffic Responsive Thresholds

Offset Change	CC Enter Value	CC Exit Value
1 to 2	25	18
2 to 3	52	49
3 to 4	68	64
4 to 5	75	70

Having separate entry and exit thresholds prevents the system from fluctuating frequently into and out of timing plans. This limitation is sometimes referred to as a hysteresis, or the delay between change in state (detector data) and the system response (change to signal timing plan). The separate thresholds in a threshold-based system provide the same type of operation as percent improvement targets in a target-based system

Some threshold-based systems calculate ratios of detector data for the selection of cycle length, splits, and offsets, referred to as "indices." A few general principles apply to this method of combining CC parameter information:

- The cycle length index is a function of the arterial volume (V+KO). The cycle length increases as the volume detected on both the inbound and outbound detectors increases.
- The split index is proportional to the ratio of arterial volume (V+KO) to crossing street volume (V+KO).
- The offset index is proportional to the ratio of inbound to outbound volume (V+KO).

9.3.1.3 Traffic Responsive Process Example

There are several steps required to set up and configure a traffic responsive system. This section walks through an example using a target-based system. The same principles can be applied to threshold-based systems; the practitioner will simply need to define the "into" and "out of" thresholds.

The first step when setting up a traffic responsive system is to identify the group of signals that will operate together. Typically, the group of signals is composed of intersections along an arterial that have substantial fluctuations in traffic volumes, requiring a range of cycle times. For this example, it will be assumed that the signalized intersections shown in Exhibit 9-3 experience variable traffic flow.

The next step in the setup process is to establish target volumes for each detector that are associated with specific types of cycle/split/offset combinations. Traffic data should be collected for several days at the input and output detector locations (and summarized in daily charts) to determine these target values. Detection stations must be installed at key input and output points. While they are not as critical, detection stations located midway along the arterial can also be configured as system detectors. The detector locations for this example are highlighted in Exhibit 9-3. (Note that the exhibit is oriented so north is at the top of the figure.)

Exhibit 9-3 Example Traffic Responsive Detector Locations

Exhibit 9-4 provides an example of how the relationship between inbound and outbound traffic can be used to identify different types of traffic flow patterns (that will be associated with different offset values). For example, the first row of the table describes the signature for a heavy inbound flow scenario, while the third row indicates a scenario in which the level of traffic in both directions is similar and moderate. The ratios of inbound to outbound traffic can be used as the traffic responsive system targets. Once an arterial moves from balanced traffic flow (with inbound and outbound traffic nearly equal) to a moderate inbound traffic flow (with inbound traffic 1.5 times that of outbound traffic), the system is notified that it is time to select a new timing plan (with different offsets).

Exhibit 9-4 Example Offset Targets (Based on Arterial Traffic Flow)

Type of Offsets (Based on Traffic Flow)	Ratio of Traffic Volume Detected at Inbound Detector to Traffic Volume Detected at Outbound Detector (Vehicles:Vehicles)		
	Inbound Detector 1: Outbound Detector 1	*Inbound Detector 2: Outbound Detector 2*	*Inbound Detector 3: Outbound Detector 3*
Heavy Inbound	2:1	2:1	2:1
Moderate Inbound	1.5:1	1.5:1	1.5:1
Balanced	1:1	1:1	1:1
Moderate Outbound	1:1.5	1:1.5	1:1.5
Heavy Outbound	1:2	1:2	1:2

The cycle length and splits that are selected will often be based on traffic characteristics at the critical intersection (in this example, Detector 2 detects the critical intersection major street volumes and Detector 4 detects the minor street volumes). Example target values for major and minor street traffic volumes are illustrated in Exhibit 9-5. In this example, if both the major street and minor street detectors indicate low, approximately equal traffic volumes, a timing plan with a shorter cycle length and balanced splits will be selected. If both the major street and minor street experience high traffic volumes, but the volume on the major street exceeds that of the minor street, then a timing plan with a longer cycle length and splits that favor the major street will be selected.

Exhibit 9-5 Example Cycle/Splits Targets (Based on Critical Intersection Volumes)

Type of Cycle and Splits	Traffic Volume Detected at Critical Intersection (Vehicles)	
	Major Street Volume (Inbound and Outbound)	*Minor Street Volume (Northbound and Southbound)*
Short Cycle/Directional Splits	800	400
Short Cycle/Balanced Splits	800	800
Medium Cycle/Directional Splits	1600	800
Medium Cycle/Balanced Splits	1600	1600
Long Cycle/Directional Splits	3200	1600
Long Cycle/Balanced Splits	3200	3200

These two matrices are merged (as shown in Exhibit 9-6) to correlate traffic patterns with specific timing plans. In other words, a specific combination of cycle/splits/offsets is assigned a pattern number. Each pattern is defined by the target volumes at each system detector. Example pattern signatures are illustrated in Exhibit 9-7 for Patterns 1, 8, 19, 22, and 30.

Exhibit 9-6 Example Pattern Assignments Based on Types of Traffic Flow and Timing Plans

Type of Cycle and Splits	Type of Offsets				
	Heavy Inbound	*Moderate Inbound*	*Balanced*	*Moderate Outbound*	*Heavy Outbound*
Short Cycle/Directional Splits	Pattern 1	Pattern 2	Pattern 3	Pattern 4	Pattern 5
Short Cycle/Balanced Splits	Pattern 6	Pattern 7	Pattern 8	Pattern 9	Pattern 10
Medium Cycle/Directional Splits	Pattern 11	Pattern 12	Pattern 13	Pattern 14	Pattern 15
Medium Cycle/Balanced Splits	Pattern 16	Pattern 17	Pattern 18	Pattern 19	Pattern 20
Long Cycle/Directional Splits	Pattern 21	Pattern 22	Pattern 23	Pattern 24	Pattern 25
Long Cycle/Balanced Splits	Pattern 26	Pattern 27	Pattern 28	Pattern 29	Pattern 30

As a final step in the configuration process, these signatures are entered into the signal system module, and the algorithm then selects the pattern that most closely matches the real-time, smoothed detection data from the system detectors. Most practitioners have had better success with traffic responsive systems when they defined a reasonable number of intersections and timing plans. Care must be taken to set up parameters appropriately and ensure that detection systems are functional. With proper calibration and set up, traffic responsive operations can provide significant reductions in travel time and delay by better matching the signal timing to the underlying traffic patterns.

Pattern Information		Pattern Signatures				
		Pattern 1	*Pattern 8*	*Pattern 19*	*Pattern 22*	*Pattern 30*
Type of Pattern		Heavy Inbound Volume Requiring Short Cycle/ Directional Splits	Balanced Volumes Requiring Short Cycle/ Balanced Splits	Moderate Outbound Volume Requiring Medium Cycle /Balanced Splits	Moderate Inbound Volume Requiring Long Cycle/ Directional Splits	Heavy Outbound Volume Requiring Long Cycle/ Balanced Splits
Ratio of Inbound Volume to Outbound Volume on Major Street		2:1	1:1	1:1.5	1.5:1	1:2
Critical Intersection Major Street Volume (Inbound and Outbound)		800	800	1600	3200	3200
Critical Intersection Minor Street Volume (Northbound and Southbound)		400	800	1600	1600	3200
Detector 1 Targets	Inbound Volume	530	400	640	1920	1070
	Outbound Volume	270	400	960	1280	2130
Detector 2 Targets (Critical Intersection Major Street)	Inbound Volume	530	400	640	1920	1070
	Outbound Volume	270	400	960	1280	2130
Detector 3 Targets	Inbound Volume	530	400	640	1920	1070
	Outbound Volume	270	400	960	1280	2130
Detector 4 Targets (Critical Intersection Minor Street)	Northbound Volume	200	400	800	800	1600
	Southbound Volume	200	400	800	800	1600

Exhibit 9-7 Example Pattern Signatures

9.3.2 Traffic Responsive Detection

Traffic responsive algorithms rely on detector data from "system" detectors to determine which timing plan or pattern to implement. Detectors must be located reasonably upstream from the stop bar and should be very short (typically zones measuring 6 feet by 6 feet) to reduce the effect of traffic queues. Detectors are frequently located on the exit side of an intersection and are separated by lane in order to yield a more accurate assessment of traffic conditions. Tube counters or other temporary detection equipment may be used to determine configuration parameters before permanent system detection stations are installed. Knowledge of local conditions is beneficial when determining detection station placement.

While functioning detectors are crucial to the effectiveness of a traffic responsive system, many systems allow plan selection calculations to be completed even when some of the detection stations are not functional or are not reporting valid data. These failed detector outputs either yield a measurement of zero, or are simply ignored, so the calculations are not affected. If too many detectors fail, however, the traffic responsive system will be disabled.

Too many transitions between timing plans can have an adverse effect on operations.

Traffic responsive methods typically smooth detector data through time-based weighted averages in order to prevent an excessive number of timing plan changes caused by short-term surges or reduced traffic volumes. Smoothing the detection data reduces the ability of algorithms to respond quickly to surges in traffic, so the application of traffic responsive systems is most effective when changes in traffic flow are both large and persistent (e.g., 30 minutes or more). A smoothing factor of 50 percent is commonly identified as a reasonable setting.

9.4 ADAPTIVE SIGNAL CONTROL TECHNOLOGY SYSTEMS

ASCT is a signal system technology that uses detection data and algorithms to adjust signal timing parameters for current conditions. Unlike traffic responsive plan selection systems, which use predetermined timing plans, ASCT is able to adjust various timing parameters (within certain constraints) based on what the traffic requires. Over time, traffic evolves—whether quickly in minutes or slowly in days, months, or years—and renders signal timing plans increasingly less effective. Traditional signal operations are often unable to adjust to this variability, resulting in degradation of traffic performance. Traffic responsive systems may require a large library of plans to address highly variable traffic conditions. Conceptually, ASCT's continual adjustment of signal timing parameters provides incremental benefits over time-of-day plans.

It is important to understand that ASCT systems are not "set-and-forget" systems. They require ongoing fine-tuning and higher levels of maintenance than traditional systems, in order to keep the detection and communications infrastructure working at a high level of performance. While ASCT is beneficial in certain applications, there are many other sound traffic engineering principles included in this manual that may improve traffic operations without the expense and complexity associated with ASCT. This section provides an overview of agency objectives that are often applied to ASCT systems, as well as operational characteristics and various factors worth considering during the selection and deployment process. For more information on agency experiences with ASCT systems, refer to *NCHRP Synthesis 403* (*4*).

9.4.1 Agency Objectives

It is critical to acknowledge that ASCT cannot solve underlying system capacity issues.

A wide range of operational objectives serve as motivation for ASCT deployment, but variable traffic demand is the most common characteristic that agencies want to address through an adaptive system. While advanced coordination and traffic responsive systems can accommodate predictable variability, ASCT systems are able to accommodate less predictable variations in traffic, which may be caused by traffic incidents, special events, or short- or long-term traffic flow changes. ASCT systems apply solutions according to specified objectives embedded in the program. The degree to which an agency is satisfied with ASCT operations often depends on how closely the ASCT's optimization objectives match the agency's objectives. Common high-level objectives for the deployment of ASCT are summarized in Exhibit 9-8.

It is often the case that different objectives are appropriate at different times of the day and under different traffic conditions. An arterial road that provides access between a freeway and large residential areas but also contains traffic generators (e.g., retail centers or schools) may require a peak-period pipeline objective during most times of the day but a smooth flow or access equity objective during business hours and on weekends. Under these conditions, the ASCT may be required to accommodate the transition of objectives at different times of day. Most ASCT systems today ***do not*** explicitly include features or configurations to address this, although some systems may be modified to transition between objectives with the detection of user-specified field conditions.

Exhibit 9-8 ASCT System Objectives

Objective	Description	Where Objective Is Typically Applied	How Objective Is Applied
Pipeline	Minimizing the number of stops experienced by a preferred movement on a critical route (from one end of the corridor to the other).	Linear, arterial routes. Less commonly applied to progress heavy turning movements.	Maintaining large splits for the coordinated phases.
Smooth Flow	Prioritizing simultaneous bidirectional movements on a critical route in order to maximize throughput.	Suburban arterials.	Maintenance of large splits for the coordinated phases.
Equitable Access	Providing sufficient arterial access to traffic generators along a corridor by placing increased emphasis on minor street demand.	Areas with significant left-turn and minor street demand (e.g., suburban retail shopping districts).	Appropriate application of split times to prevent long delays for minor movements (including pedestrians).
Manage Queues	Mitigation of queues and congestion caused by blocked intersections or movements. *Note: Most ASCT systems do not include features that specifically mitigate queues.*	Locations where queues block upstream intersections or movements.	Constraints on cycle and phase durations to ensure that large platoons progress at appropriate times. “Gating” to store vehicles at locations upstream of critical links may also be required.
Mitigate Oversaturation	Prevent, delay the onset of, or limit the duration of oversaturated conditions. If oversaturated conditions do persist, clear overflow queues quickly.	Locations with oversaturated movements. *Note: ASCT systems can mitigate some oversaturation due to short-term capacity limitations but are not a substitute for capacity improvements.*	Adjusting green time allocation for saturated phases.
Accommodate Long-Term Variability	Update signal timing more frequently than traditional systems and reduce deterioration of traffic operations over time.	Areas with changing traffic patterns, particularly growing communities.	While ASCT systems require some adjustments when major traffic pattern changes occur, they can adjust automatically to many changes.
Manage Events and Incidents	Manage surges in traffic (both planned and unplanned).	Areas with recurring planned special events (e.g., concerts, sporting events, or community activities).	Adjust signal timing parameters to match traffic surges over time. However, they generally cannot adjust rapidly to large surges.

9.4.2 Operational Characteristics

ASCT systems adjust traffic signal settings based on current traffic conditions (or nearly current, as most systems lag changes in demand by one or more cycles). However, available systems vary widely in levels of responsiveness, algorithmic framework, and detection requirements. The practitioner should refer to the manufacturer for specific information about each system. This section summarizes a few operational characteristics that are common across many systems. In particular, systems share the same basic structure:

- Detector collection of traffic conditions information,
- Calculation of timing plans based on inputs from detection, and
- Selection of the next settings to be applied.

9.4.2.1 Adaptive Groups

Adaptive systems often require the division of an ASCT-controlled network into subsystems (or regions) of intersections that either need to be coordinated or should operate using a common cycle length. This grouping is determined during the design phase. Typical rules for grouping intersections using traditional signal timing also apply to the identification of groupings for adaptive operations (refer to Chapter 3).

Some ASCT systems allow for "cross-coordination" on multiple, crossing routes, while others only allow coordination along a single route (e.g., north/south, east/west) by time of day. Some ASCT systems allow a bordering intersection in one subsystem to leave its current subsystem and join a neighboring subsystem according to a suitability factor calculation. These types of settings need to be calibrated during system installation and configuration.

9.4.2.2 Calculation Methods

ASCT uses two primary methods to adjust signal timing: (1) downloading new settings and (2) overriding the local controller. Some ASCT systems compute new signal timing parameters and download those new settings to the local controller. ASCT systems that download new parameters allow the local controller to use actuated operational features to adjust green splits according to gap out and fixed or floating force-off logic. Alternatively, ASCT systems can override the operation of the local controller through hold, force-off commands, or through controlling the presence or absence of phase calls. Through override methods, some ASCT systems use command messages to communicate the changes to the controller, while others interface to the cabinet directly using either hardwire electrical connections (for Type 332, Type 336, and NEMA TS-1 cabinets) or serial interface units (for NEMA TS-2 and ITS cabinets).

ASCT systems that download new timing parameters are typically unconstrained by algorithm processing time, but limit the breadth of their search for better timing to a limited range around the current settings. This is done for two reasons: (1) to limit the effects of signal transition from the current settings to the new settings and (2) to minimize "over-reaction" to very short-term anomalies in flows. For example, if a current split value is 15 seconds, an ASCT optimization method might consider values from 10 seconds to 20 seconds for the next split value. If the current cycle time is 100 seconds, cycles between 110 seconds and 120 seconds might be considered next (even if a better value were 130 seconds).

Some ASCT systems are not as concerned about minimizing the difference between the last phase split and the next phase split because they have complete control of phase durations, and no transitions are necessary. This is sometimes referred to as a rolling-horizon optimization process. These systems constantly re-optimize the next signal settings in reaction to short-term fluctuations by trying to predict what the best possible sequence and duration of phases should be in the next few seconds to minutes. For example, a rolling-horizon system might evaluate hundreds of different options for the duration and sequence of each phase over the next 60 seconds and then implement only the decision trajectory for the next 5 seconds. The process then repeats. Maximum and minimum phase durations are typically enforced much like a controller running in free mode.

All ASCT systems have settings that need to be fine-tuned to customize the level of adaptability. These settings can include maximum and minimum green times for each phase, adherence to pedestrian crossing time requirements (or the ability to ignore pedestrian times if there is no pedestrian demand), the amount each timing parameter may be changed incrementally, the largest and smallest values for the timing parameters (e.g., cycle lengths or offsets), and the timing parameters that the system is prohibited from changing (e.g., phase sequence, ability to skip phases).

The amount of time required for calibrating and fine-tuning models and adaptive operations is directly related to the number of signals that the ASCT manages.

9.4.2.3 Detection

Detection installation, configuration, and accuracy are ***critical*** to the effective operation of ASCT because the results are used by the optimization algorithms to determine appropriate signal timing settings. However, ASCT systems are able to receive input from various detection technologies (e.g., microwave, magnetometer, inductive loop, and video). Some systems are even integrated with specific types of detection devices.

ASCT systems often require extensive detection that must be highly reliable and well maintained.

Many ASCT systems may have a preferred (and some a required) detection configuration, but can operate a subset of adaptive algorithms with a different configuration. Generally, there are five locations where ASCT detectors ***may*** be needed:

- Stop bar detectors,
- Setback detectors (typically used for decision zone protection and phase extension in actuated operations),
- Mid-block detectors,
- Upstream detectors located at the upstream intersection (on the downstream side), and
- Setback detectors at the entrance points of left- and right-turn storage bays.

No ASCT system requires detection at all five locations, but detection is generally required on all phases. While many ASCT systems often require lane-by-lane detection on all phases (which provides the best ASCT operations), some are able to accommodate detection that is reported across multiple lanes for each phase.

Typically, stop bar detectors estimate the degree of saturation for each phase and measure output for queue length estimation. These detectors often need to be added to coordinated phase lanes if the existing operation is coordinated but not actuating the coordinated phase(s). Shorter stop bar detection zones yield more accurate count and occupancy information; although in practice, newly deployed ASCT systems often accommodate existing stop bar detection zones that are 50 feet or longer.

Setback detectors can estimate queue lengths and measure the arrival profile of approaching platoons of traffic. The farther upstream these detectors are located, the better the quality of the arrival profile information. It is worth mentioning that mid-block arrival flows entering the link downstream of the detection point will not be captured by the ASCT, so care must be taken to locate setback detection appropriately on a link-by-link basis. It is also important to locate setback detection away from driveways so that vehicles exiting the street do not introduce error into the ASCT calculations.

9.4.2.4 Communications

Communication among the detectors, controllers, and ASCT system is vital. Some ASCT systems interact with detection directly, either (1) through direct hardwire integration with the devices or (2) through wireless polling of stand-alone detection. Others rely on communication with the local controller to retrieve detection status. Many systems can accommodate a combination of directly connecting to detection and receiving detection status from the local controller. A detailed understanding of existing communications capabilities and ASCT needs is essential when evaluating alternative systems and their overall cost. Communications infrastructure has been a challenging requirement for ASCT in the past, although the emergence of Internet Protocol (IP) technology in traffic operations has improved the deployment options. (IP communications can either be achieved through wireless technology or through hardwire technology.)

9.4.2.5 Phase Configuration

The practitioner should be aware of the phase configuration limitations of the ASCT system. Some systems can currently only operate with eight-phase, dual-ring controller architecture, while other technologies can accommodate more than eight phases. Many ASCT systems also use stage-based operation instead of phase-based operation, requiring configuration of all allowable combinations of phases as stages (e.g., phase pairs 1 and 6, 1 and 5, 2 and 5, 2 and 6, etc. are configured as separate stages).

9.4.3 Installation and Configuration

The installation of an ASCT system can be, but does not have to be, a lengthy and difficult process. The condition and suitability of the existing infrastructure (i.e., detection, hardware, and communications) are the most critical elements of the deployment. Common activities during installation and configuration of an ASCT system are described below.

9.4.3.1 ASCT Equipment Installation

- Installation of any special ASCT equipment in existing traffic cabinets or construction and installation of new cabinets as needed.
- Installation and configuration of IT systems, databases, servers, and computer peripherals for ASCT operation.

9.4.3.2 ASCT System Configuration

- Installation and configuration of a plan and schedule that is both realistic and tolerant of technology-related issues.
- Configuration and calibration of ASCT adaptive and model parameters.
- Identification and configuration of adaptive on and off times, if desired.

- Configuration of interaction with third-party systems such as preemption, transit priority, and pedestrian countdown timers.

9.4.3.3 Detection Installation

- Installation of new detection or reconfiguration of existing hardware by qualified contractors.
- Separation of detection lead-in cables into lane-by-lane operations.
- Addition of detector cards, detector racks, video system processors, and other detection equipment.
- Verification of all detection operations.

9.4.3.4 Communications Installation

- Installation of new communications or upgrade of existing communications media between traffic signals and from signals in the field to a central or master location.
- Configuration of communications equipment (i.e., modems, switches, firewalls, and central server computers).
- Verification of all communications operations.

9.4.4 Maintenance and Operations

ASCT systems are more operationally demanding and require more agency support than traditional traffic signal systems. As discussed in *NCHRP Synthesis 403*, ASCT systems rely on experienced staff that understand and can adjust operations (*4*). The need for training and ongoing operational attention of agency staff needs to be recognized in the early stages of ASCT consideration. One solution for an agency may be to obtain operational support from outside sources, including engineering consultants or system vendors. Whether the adaptive system is maintained through contracted support, development of existing staff, or newly hired talent, it is critical that agencies never consider an ASCT system to be a "set-and-forget" system that eliminates the need for sound traffic engineering and expertise. Key activities during the maintenance and operations phase of ASCT deployment are described below.

9.4.4.1 ASCT System Review

- Review and analysis of signal timing policies and underlying signal settings.
- Frequent review of ASCT decision-making for recalibration.
- Frequent analysis of MOEs for adaptive parameter fine-tuning.
- Development of backup timing plans.

9.4.4.2 Equipment Maintenance

- Continual preventative maintenance checks on field hardware, communications, and detection systems.
- Timely repair of malfunctioning detection and communications hardware.

- Identification of preferred failure modes if detection, communications, or ASCT equipment fails (e.g., coordination versus free operation).
- Database backup and software and hardware maintenance plan development.

9.4.4.3 Training and Vendor Contact

- Initial and continued training on system operations (particularly new training for staff added after ASCT deployment).
- Identification of a specific person as the ASCT vendor point of contact.

9.4.5 Verification and Validation of Effectiveness

Verification is necessary if a system is being procured with new features that have not been deployed previously in order to determine whether the agency's requirements have been met. This would not be necessary if the procurement is for an existing ASCT that has been previously verified.

Validation is the process of assessing how the deployed system actually performs relative to agency operational objectives and actual traffic conditions. The assessment of ASCT system performance is highly dependent on both the quality of timings in place before ASCT is deployed and the duration of time that occurs between the collection of before data and after data. That is, if timings are updated just before ASCT deployment, the system may experience only minor improvements or, in some situations, signal timing with actuated control may outperform the adaptive system.

The performance of adaptive systems may be evaluated through one or more performance measures. The validation process helps to ensure that these performance measures are selected and applied under the guidance of sound traffic engineering practices and principles. Further discussion on validation and MOEs may be found in *Every Day Counts: Validate Traffic Signal Operational Objectives—Draft MOE and Evaluation Approach Plan* (*5*). Some common types of validation activities include

- State operational objectives for ASCT installation clearly and measure performance with respect to those objectives; match MOEs closely with operational goals.
- Validate minor street performance, particularly in lieu of short-lived manual traffic studies that only capture a snapshot of performance. Most ASCT and controller hardware can now continuously provide detailed, high-resolution, event-based signal timing data.
- Supplement probe travel time runs with vehicle re-identification technologies, which can provide continuous measurement of point-to-point travel times. Validation activities using only probe data are limited by the number of runs and the time periods in which they are conducted.
- Evaluate ASCT effectiveness before and after peak conditions. In these time periods, ASCT effectiveness may be most dramatic, potentially revealing the ability to shorten the peak period or more quickly dissipate queues. Performance data during peak periods, particularly when conditions are oversaturated, may not indicate superiority of ASCT.
- Simulate incident conditions in order to assess how ASCT will manage abnormal conditions.

- Collect before-and-after performance data as close to the same date as possible. Separation of before-and-after data by many months can skew results. Consider an on/off approach where days are alternated with and without ASCT operation.
- Collect enough performance data for both on and off periods to estimate the reliability of performance with and without ASCT.
- Consider the sensitivity of performance due to changes in volumes. Collect volume data on key routes to identify whether conditions are similar during on and off periods.
- Consider that retiming signals just before ASCT deployment will result in less impressive ASCT validation results.

9.5 REFERENCES

1. National ITS Architecture Team. *Systems Engineering for Intelligent Transportation Systems*. Report FHWA-HOP-07-069, Federal Highway Administration, United States Department of Transportation, 2007.
2. Fehon, K., M. Krueger, J. Peters, R. Denney, P. Olson, and E. Curtis. *Model Systems Engineering Documents for Adaptive Signal Control Technology Systems*. Report FHWA-HOP-11-027, Federal Highway Administration, United States Department of Transportation, 2012.
3. Abbas, M. M., N. A. Chaudhary, A. Sharma, S. P. Venglar, and R. J. Engelbrecht. *Methodology for Determination of Optimal Traffic Responsive Plan Selection Control Parameters*. Report FHWA/TX-04/0-4421-1, Texas Department of Transportation, Austin, TX, 2004.
4. Stevanovic, A. *NCHRP Synthesis 403: Adaptive Traffic Control Systems: Domestic and Foreign State of Practice*. Transportation Research Board of the National Academies, Washington, D.C., 2010.
5. Gettman, D. *Every Day Counts: Validate Traffic Signal Operational Objectives—Draft MOE and Evaluation Approach Plan*. Report FHWA-HOP-12-034, Federal Highway Administration, United States Department of Transportation, 2012.

CHAPTER 10
PREFERENTIAL TREATMENT

CONTENTS

LIST OF EXHIBITS

CHAPTER 10. PREFERENTIAL TREATMENT

The concepts presented throughout this manual have emphasized an outcome based approach to traffic signal timing. Preferential treatment is an application that can be used at signalized intersections to adjust operations in favor of a particular user. This chapter provides an overview of the signal timing treatments that can be applied to prioritize rail, emergency vehicles, transit vehicles, and trucks.

10.1 TYPES OF PREFERENTIAL TREATMENT

Traffic signal ***preemption*** *interrupts normal operations (breaking coordination) to serve preferred vehicles, while traffic signal* ***priority*** *maintains system coordination when serving priority requests.*

The concept of preferential treatment began with the development of traffic signal preemption, initially designed as a way to immediately alter traffic signal operations near at-grade railroad crossings. Once vehicle detection improved, preemption applications eventually expanded to include emergency vehicles. Under a preemption sequence, normal operations are immediately interrupted in order to serve the preferred vehicle, without regard for the state of the signal. This can cause disruption of coordination, pedestrian service, and phasing patterns, but the severity of impacts depends on several factors, including the timing parameters, intersection spacing, transition algorithm, level of saturation, duration of preemption, and amount of slack time available in the intersection cycle (*1*).

A limited number of bus transit operations initially used preemption, but this application was largely abandoned due to the frequent disruption of signal timing and negative impacts on coordination and traffic progression. The desire for preferential treatment of transit eventually led to the development of traffic signal priority concepts. A priority sequence will maintain signal coordination while allowing timing adjustments to accommodate preferred vehicles. A current limitation of priority systems is that they are largely first-come-first-serve. Most controllers are constrained by the information they receive about future requests. While some systems can schedule requests using setback detection and communication, requests can also be managed from a fleet or transit management center. Exhibit 10-1 provides general definitions for preemption and priority applications.

Exhibit 10-1
Preferential Treatment Applications

Preferential Treatment	Definition
Traffic Signal Preemption	Per *NTCIP 1202:2005* (*2*), preemption is the transfer of the normal control (operation) of traffic signals to a special signal control mode for the purpose of serving railroad crossings, emergency vehicle passage, mass transit vehicle passage, and other special tasks, the control of which requires terminating normal traffic control to provide the service needs of the special task.
Traffic Signal Priority	Per *NTCIP 1211* (*3*), priority is the preferential treatment of one vehicle class (such as transit vehicles, emergency service vehicles, or commercial fleet vehicles) over another vehicle class at a signalized intersection without causing the traffic signal controller to drop from coordinated operations.

10.2 INTRODUCTION TO PREFERENTIAL TREATMENT

While there are various preferential treatment applications, the procedure for prioritizing a vehicle generally remains the same. There are five main steps (illustrated in Exhibit 10-2) that define the preferential treatment process:

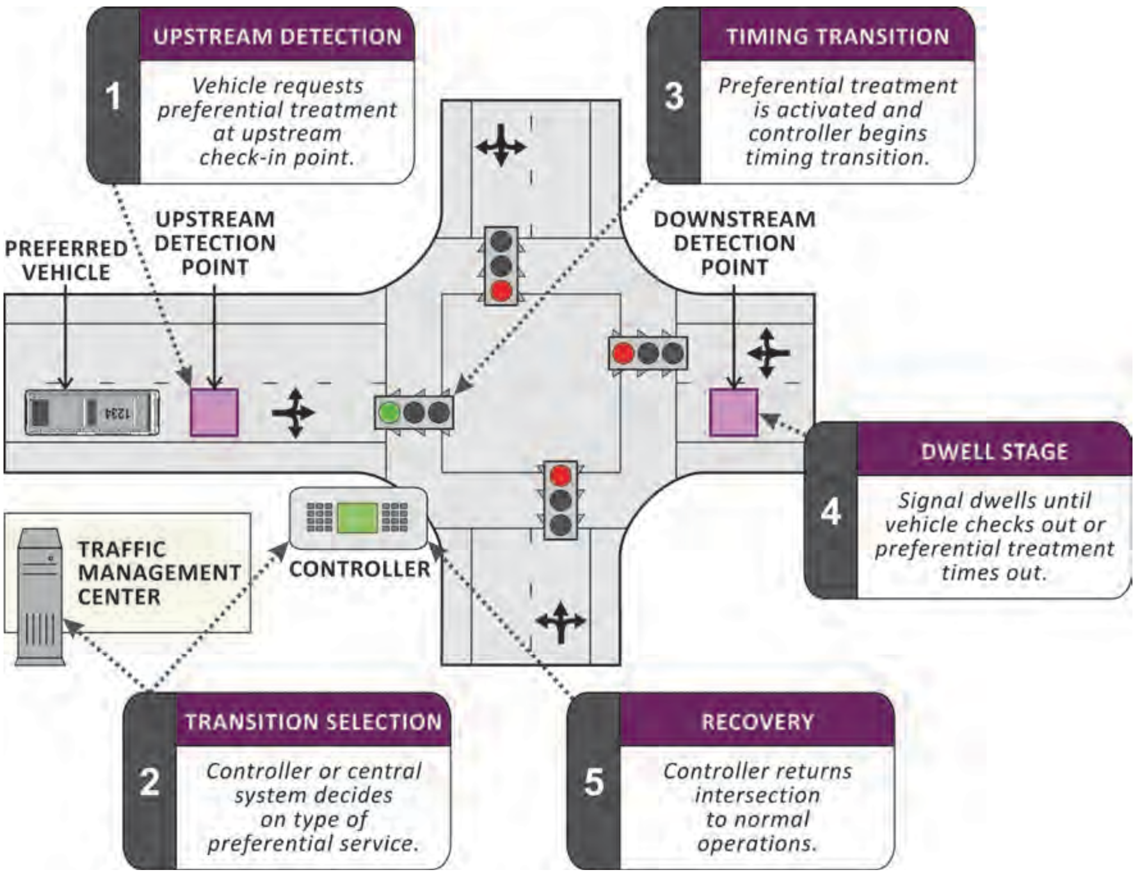


Exhibit 10-2
Preferential Treatment Process

1. ***Upstream Detection.*** The preferred vehicle sends the system a "request" for preferential treatment through an upstream detector. Depending on the type of controller firmware used in the deployment, this can be a request for immediate service (common in older controller firmware) or for the time of service desired (preferred if available). Types of preferential treatment detection are explained in Section 10.2.1.
2. ***Transition Selection.*** Depending on the status of the signal controller (as well as the status of higher priority requests), the controller or central system selects which signal timing transition to apply (e.g., green extension, red truncation). Types of transitions are described in Section 10.2.2.
3. ***Timing Transition.*** Preferential treatment is activated, and the controller begins a right-of-way transfer procedure. During this process, the signal controller safely transitions (or terminates if necessary) all phases in conflict with the signal indications requested by the preferred vehicle and transfers the right-of-way to the desired phase(s).
4. ***Dwell Stage.*** After the desired signal state is reached, the controller dwells in that state until either the preferred vehicle clears the intersection and is

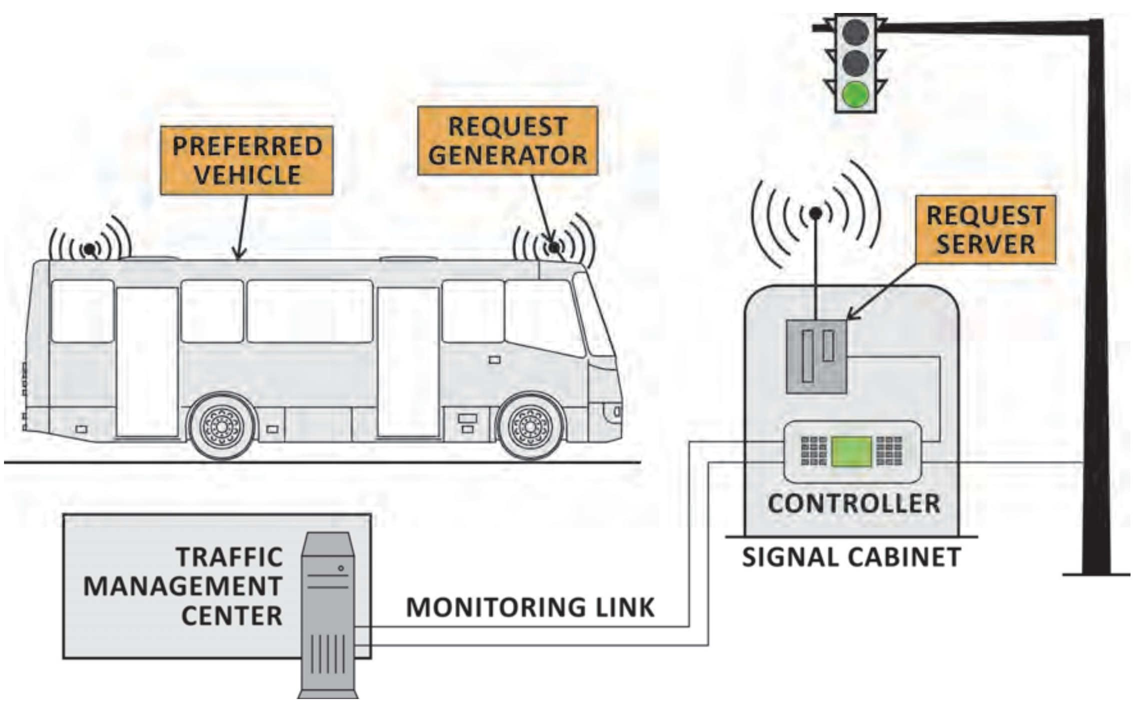
detected at a downstream detector or the preferential treatment interval times out (as programmed in the controller). Generally, limited service (in which the controller serves phases not in conflict with the preferred movement) is the preferred mode of dwell because it minimizes intersection delay.

5. ***Recovery.*** After the preferred vehicle leaves the intersection, the signal can begin the recovery stage. Recovery (exit) phases are selected by the practitioner and are usually composed of those phases that have been adversely affected by preemption operations. Once the recovery phases have been served, the signal controller reverts to normal operations through firmware-specific predetermined logic. If the controller operates in coordination, the local timer may require anywhere from one to five cycles to transition back to master clock coordination, as the preferential treatment strategy influences the time discrepancy after service. Recovery strategies are described in Section 10.2.3.

10.2.1 Detection Requirements

At a basic level, preferential treatment detection is similar to standard vehicular detection in that it requires (1) the preferred vehicle to be detected in the roadway and (2) the detection to be registered by the traffic control system. An important difference is that each preferred vehicle must be equipped with a request generator that communicates to the traffic control system. Exhibit 10-3 illustrates a preferential treatment detection system, where there is a request generator on the preferred vehicle sending a detection signal to the request server in the signal cabinet (*3*). Detection technologies differ in how preferential treatment requests are detected, and while some detection technologies register a "check-in," others emit a continuous priority request call (similar to vehicular presence detection).

Exhibit 10-3 Preferential Treatment Detection Equipment



The detection should be recognized far enough in advance so that the signal controller can conclude other competing activities before serving the request. However, detection too far in advance is not desirable due to the variability in intersection arrival times. Older firmware may only have the capability to serve preferential treatment requests immediately, while more advanced systems can use "time of service desired" (TSD) and "time of estimated departure" (TED) to better manage timing transitions and competing requests. TSD is the estimated arrival time of the preferred vehicle at the stop bar, while TED is the estimated departure time of the preferred vehicle clearing the intersection (*3*). Note that current detection inputs to the controller use a simple contact closure and not an exchange of data for TSD and TED, which means that TSD and TED ***requests*** are sent immediately.

Some of the most common types of detection used for preferential treatment are described in Exhibit 10-4 (*4*). GPS-based detection systems use the most advanced detection technology and are increasingly being implemented. However, infrared detection systems remain the most widely used.

Exhibit 10-4 Types of Preferential Treatment Detection

Detection Type	Equipment Required	Advantages	Limitations
Vehicle-Based GPS	□ In-vehicle computer that uses GPS to update vehicle location continuously □ Field unit in cabinet	□ No unobstructed line-of-sight requirement □ Notification when a vehicle has cleared the intersection □ Potentially larger detection range (i.e., requests received sooner)	□ Some systems may not adequately sample vehicle locations for accurate information at closely spaced intersections □ Acquiring satellites in urban environments can be challenging
Hard-Wired Loop	□ Transponder attached to underside of vehicle (coded with unique identification numbers for automatic vehicle identification) □ Loop of wire embedded in pavement	□ Similar to commonly used loop detectors □ No unobstructed line-of-sight requirement □ Relatively easy to implement a downstream check-out detector □ Overall very reliable	□ Requires in-pavement detectors, which need to be appropriately placed and maintained □ Distance-based, rather than time-based, requiring an estimate of speed (which may change with traffic conditions)
Infrared (Light-Based)	□ Infrared strobe emitter on the vehicle (which can contain a unique identification number for tracking purposes) □ Infrared detectors at each intersection □ Detection interface device in the cabinet	□ Widely used, allowing regions to utilize uniform systems for emergency and transit vehicles □ Technology has been well tested during its many years in use	□ Requires line of sight between the vehicle and detector; effective operation can be hindered by roadway geometry, weather problems, and obstructions such as tree foliage
Radio-Based	□ Radio frequency (RF) transponders mounted on the vehicle □ Upstream RF tag readers □ RS-232 connector to connect tag reader to signal controller	□ No unobstructed line-of-sight requirement	□ Requires suitable curbside locations for tag readers, including mounting locations, power, and communications connections

Detection Type	Equipment Required	Advantages	Limitations
Sound-Based (Siren-Based)	◻ Siren on vehicle ◻ Sound detector at intersection	◻ No new vehicle equipment	◻ Less precise control of activation
Push Button	◻ Push button (e.g., in fire station)	◻ Allows activation before vehicle departs	◻ Requires manual activation
Track Circuits	◻ Designed by railroad or transit agency	◻ Advance, simultaneous, and gate-down notifications provide maximum flexibility of operation	◻ Upgrades from simultaneous-only can be very expensive

10.2.2 Signal Timing Strategies

How requests for service are accommodated at a particular intersection depends on many variables, including (but not limited to):

- Controller firmware capabilities,
- Agency policy,
- Cycle length,
- Complexity of phases,
- Traffic on intersecting streets,
- Protection of minimum clearance times for pedestrians,
- Minimum phase times, and
- Accuracy of check-out mechanisms.

Regardless of the traffic signal controller's strategy for providing preferential treatment, it is important to have an understanding of the parameters and expected outcomes. This section explains the various signal timing methods that can be used to accommodate preferential treatment. Some strategies are used with priority and can maintain coordination (e.g., green extension and phase skipping), while others are used with preemption and are more disruptive, typically requiring recovery strategies to return to coordinated operations. It should be noted that if preemption is frequent, transitions back to coordination may result in frequent disruptions, and uncoordinated operations should be considered.

10.2.2.1 Green/Phase Extension

Green (or phase) extension is a common strategy used to serve preferential treatment requests. Green extension involves the extension (or holding) of the preferred phase green interval past its normal termination point (as illustrated in Exhibit 10-5). Depending on the type of system and detection scheme, the extension period can be for a fixed duration or until the preferred vehicle has cleared the intersection (subject to a maximum extension). This strategy is designed to prevent long delays for preferred vehicles that are anticipated to arrive near the end of the green interval.

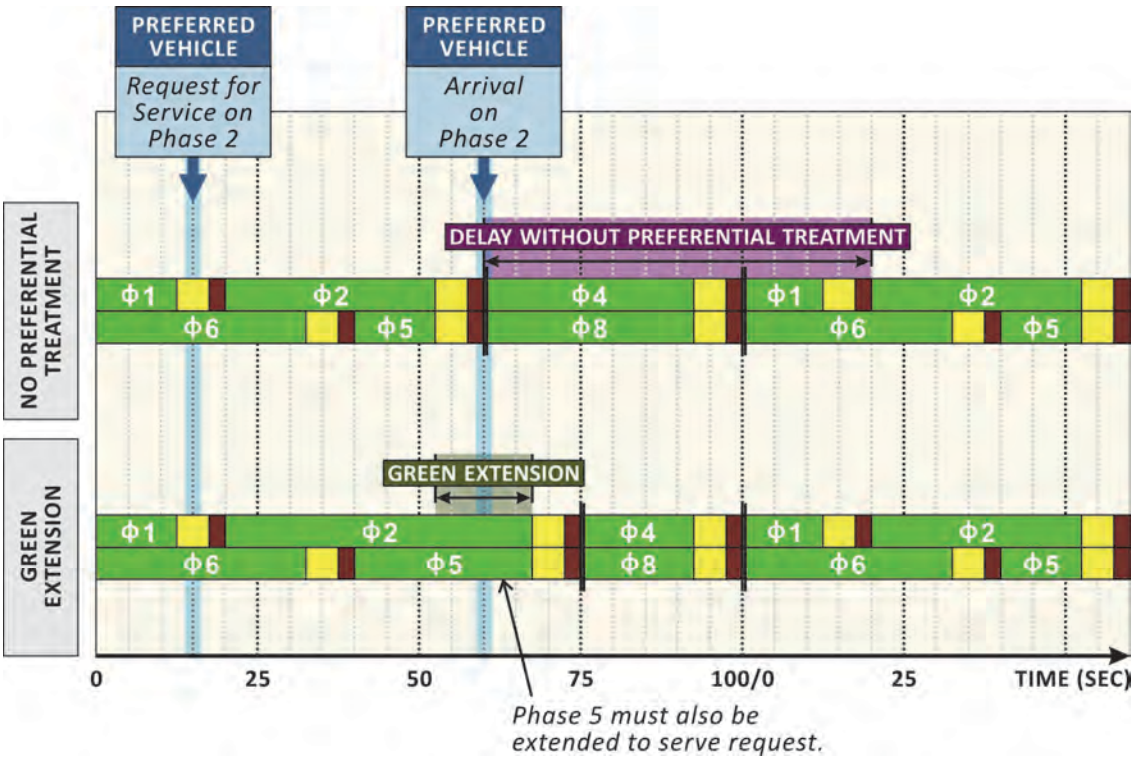


Exhibit 10-5 Green/Phase Extension

10.2.2.2 Red Truncation/Early Green

Red truncation (or early green) is a preferential treatment strategy that shortens the duration of non-preferred phases in order to return earlier than normal to the green interval of the preferred phase (as illustrated in Exhibit 10-6).

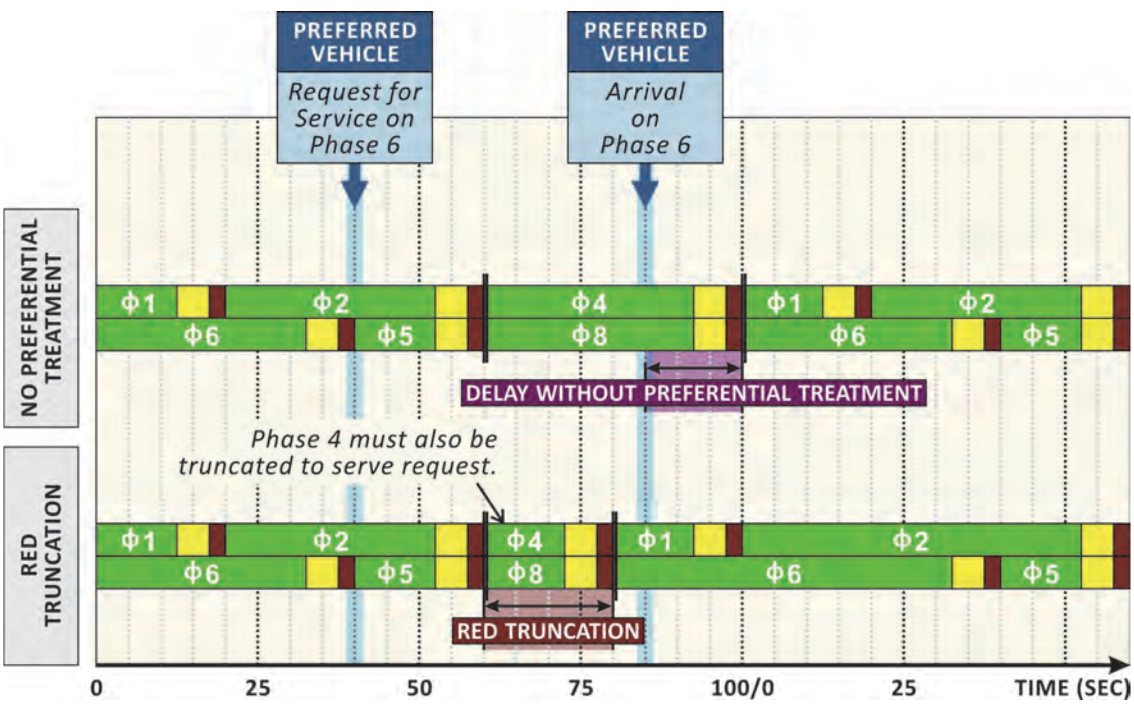


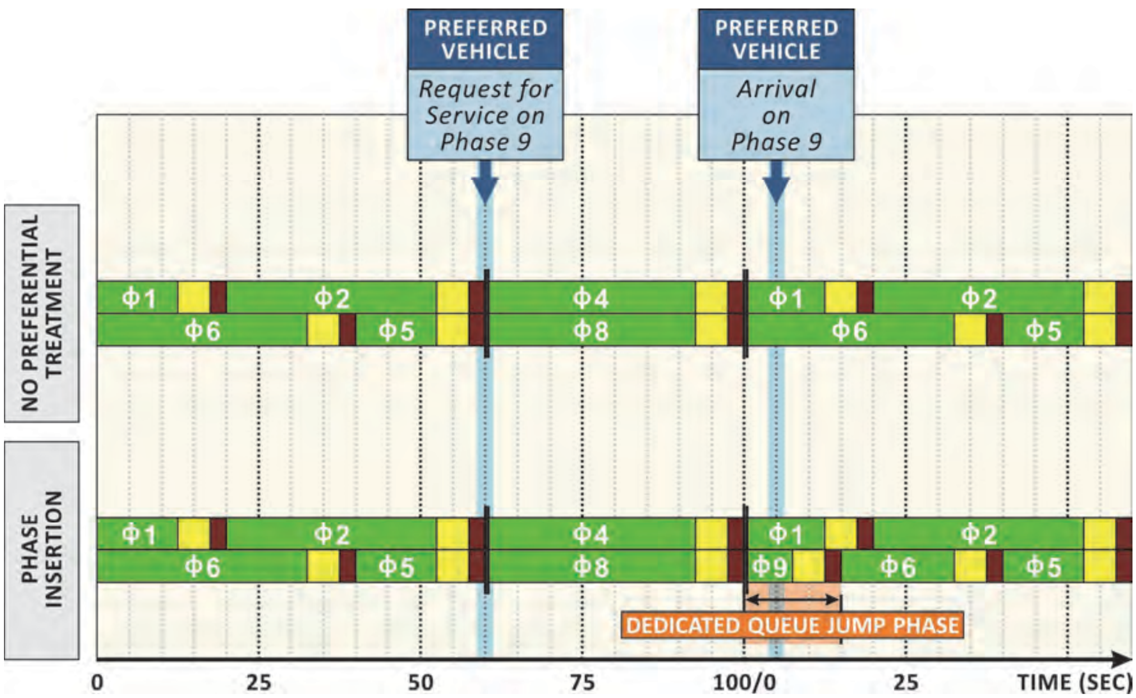
Exhibit 10-6 Red Truncation/Early Green

In this strategy, the duration of some or all of the non-preferred phases can be reduced. How quickly the traffic signal controller can return to the preferred phase is constrained by the minimum green time and clearance requirements of the non-preferred vehicular and pedestrian phases. This strategy is less beneficial than green extension and may be undesirable at locations with competing preferred vehicles.

10.2.2.3 Phase Insertion

Phase insertion involves the activation of a special, dedicated phase that is not served during normal (non-preferred) operations and is only displayed when a preferred vehicle has been detected at the intersection. This strategy is commonly used to provide service to lanes that are dedicated to preferred vehicles only (e.g., an exclusive left-turn lane into a transit transfer point). This strategy is also used to support queue jumps that allow preferred vehicles to enter a downstream link ahead of the normal traffic stream. Exhibit 10-7 shows the use of phase insertion with a queue jump strategy. Note that the dedicated queue jump phase is Phase 9. Depending on whether the controller can accommodate more than eight phases, this phase insertion strategy may need to be managed differently (e.g., unused phase, if available) than shown in Exhibit 10-7.

Exhibit 10-7 Phase Insertion



10.2.2.4 Sequence Change

With sequence change, the order of the signal phases is altered to provide more immediate service to the preferred vehicle (as shown in Exhibit 10-8). In this example, changing a left-turn phase from leading to lagging reduces the wait time of the preferred vehicle. However, the practitioner should ensure that a sequence change will not cause operational issues; the phase sequence considerations from normal (non-preferred) operations should not be ignored (e.g., if a lead-lag sequence was chosen to accommodate intersection geometry). See Chapter 5 for guidance on phase sequence.

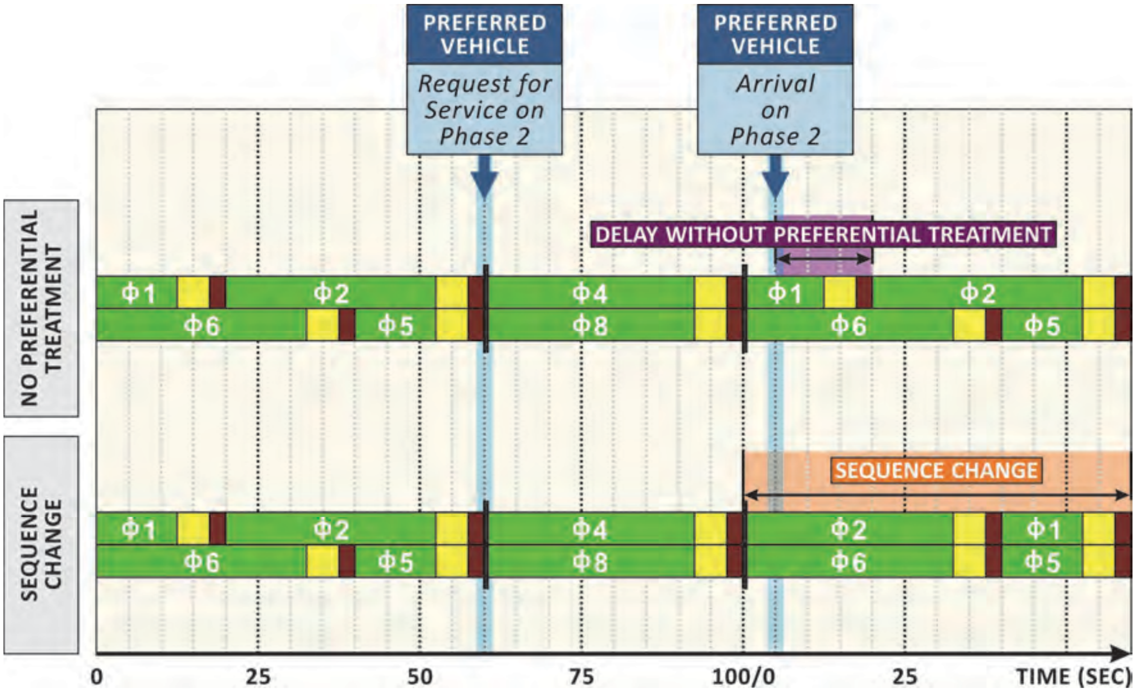


Exhibit 10-8 Sequence Change

10.2.2.5 Phase Skipping

Phase skipping (shown in Exhibit 10-9) forgoes service to (or skips) non-preferred phases that would normally be served, in order to serve the preferred phase more quickly. For example, phase skipping may skip the protected interval at a protected-permitted left-turn movement when needed. Because of the potential impact this strategy can have on delays to non-preferred movements, the practitioner should consider tradeoffs carefully when implementing this strategy.

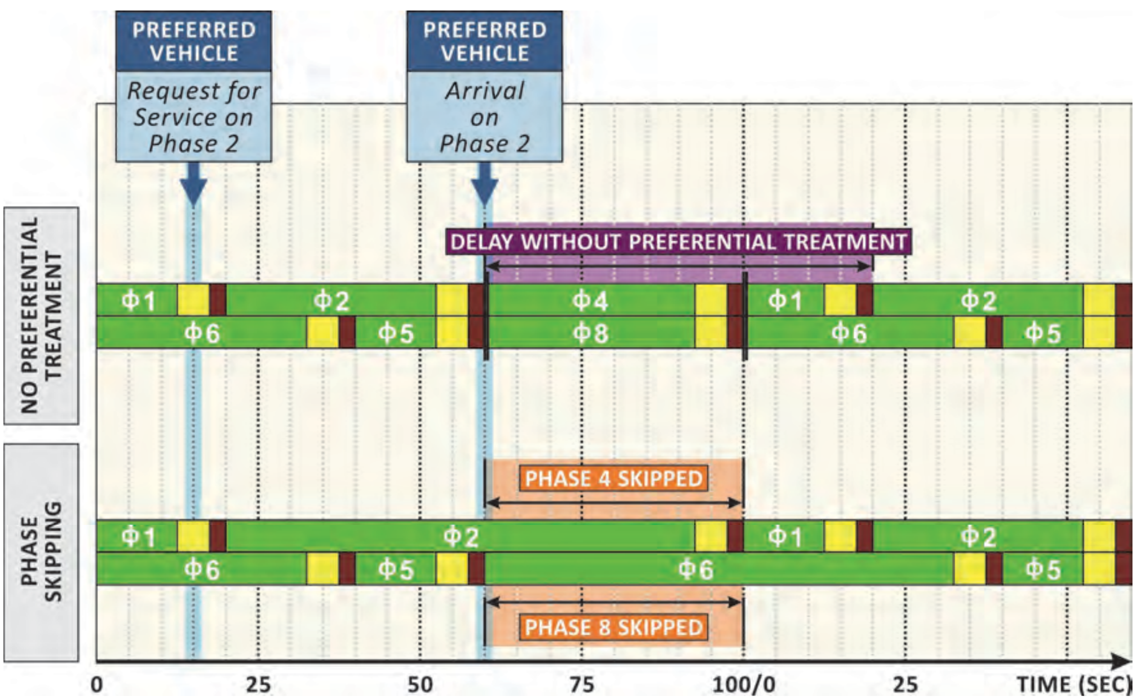
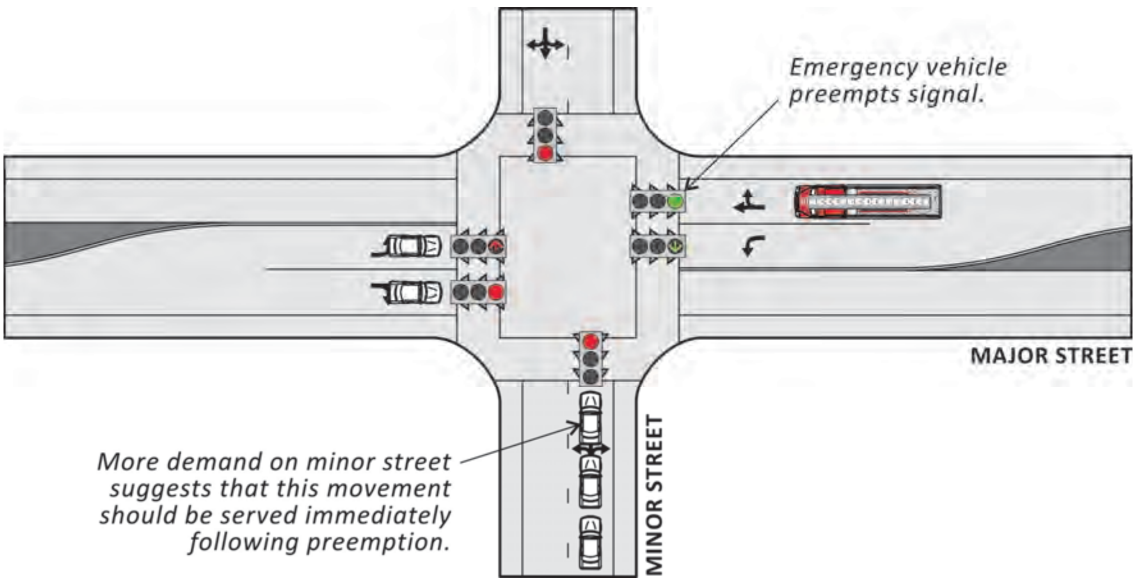


Exhibit 10-9 Phase Skipping

10.2.3 Strategic Recovery

To maintain operations at an intersection, a practitioner should consider a post-preferential-treatment "recovery" (or "exit") strategy. While preferential treatment can help meet specific objectives by prioritizing certain users, there are tradeoffs for the overall intersection operations. A clear understanding of the desired outcomes and tradeoffs of preferential treatment is necessary when selecting a recovery strategy. For example, Exhibit 10-10 illustrates an intersection under typical emergency vehicle preemption. If the minimization of intersection delay and queuing is the primary operational objective, then the most appropriate recovery strategy could be to serve the minor street approach before returning to normal operations.

Exhibit 10-10 Example of Post-Preferential-Treatment Recovery



A recovery strategy can help mitigate the effects of preferential treatment for non-preferred users, particularly when

- ***Preferential treatment has a long duration***, leading to long delays for movements that are prohibited during preferential treatment.
- ***Preferential treatment requests are frequent*** (i.e., back-to-back or near back-to-back), leading to substantially shortened or skipped phases or movements, such that long, disproportionate queues and delays occur.

There are a variety of strategies that can be applied at an intersection after preferential treatment. These recovery options are increasingly becoming available within traffic signal controller firmware. Those that can dynamically respond to prevailing conditions are the most advanced, but a dynamic recovery strategy is not necessary for all locations or time periods. Recovery strategies include

- ***Return to Normal Operations.*** Return to normal traffic signal timing; this is often the default.
- ***Return to Free Operations.*** Return to free operations for a defined period of time; this allows the signal to leverage longer green intervals and clear vehicular queues.

- ***Return to Coordinated Operations.*** Tracks coordinated timing parameters in the background for immediate return to coordination, without the need for a transition period to correct offsets.
- ***Return to Alternate Plan.*** Return to coordinated operations through the use of alternate timing parameters (e.g., cycle, splits, offsets, and/or phase order) for a defined period of time, in order to clear queues.
- ***Return to Interrupted Phase(s) (Priority Return).*** Return to vehicular and/or pedestrian phases that were interrupted at the onset of the preferential treatment request.
- ***Return to Defined Phase(s).*** Invariable return to defined vehicular and/or pedestrian phase(s) upon exit of preferential treatment.
- ***Queue Delay Recovery.*** Dynamic return to movements that have the highest delay, volume, or combination of both (based on detection); most effective where frequent priority calls occur and if coordination does not necessitate immediate restoration.

10.2.4 Data Logging

Much like any signal timing, monitoring and maintenance of preferential treatment will help the system continue to operate as desired. Controller logs and centralized recording software are available to help practitioners log and better understand preferential treatment operations. Data logs can provide information about when, where, how, and who requested preferential treatment so that preferential treatment rights and signal timing settings may be actively managed.

Logs and central tracking software are important to the active management of an effective preferential treatment system.

10.3 PREFERENTIAL TREATMENT ADVANCEMENTS

The historical approach to preferential treatment has been to serve only one request at a time. However, traffic signal controller firmware is increasingly featuring advanced preferential treatment options that allow practitioners to assign levels of priority to users (*1*). For example, a bus can be assigned a higher level of priority than a truck. This approach overcomes many of the limitations of traditional preferential treatment by scheduling calls based on the assigned priorities.

With a schedule-based approach, the controller firmware assigns priorities to requests as they are received. The highest priority request is always served at its scheduled time. If possible, the firmware then continues to serve other requests in order of priority. Lower priority requests that interfere with higher priority requests must yield to the higher priority users. While schedule-based preferential treatment does not eliminate conflicting requests, the treatment facilitates the resolution of conflicts based on priorities. In many cases, an immediate response to a call is unnecessary, and efficiencies may be realized with the application of preferential treatment scheduling (*5, 6*). Other benefits may include (*5*)

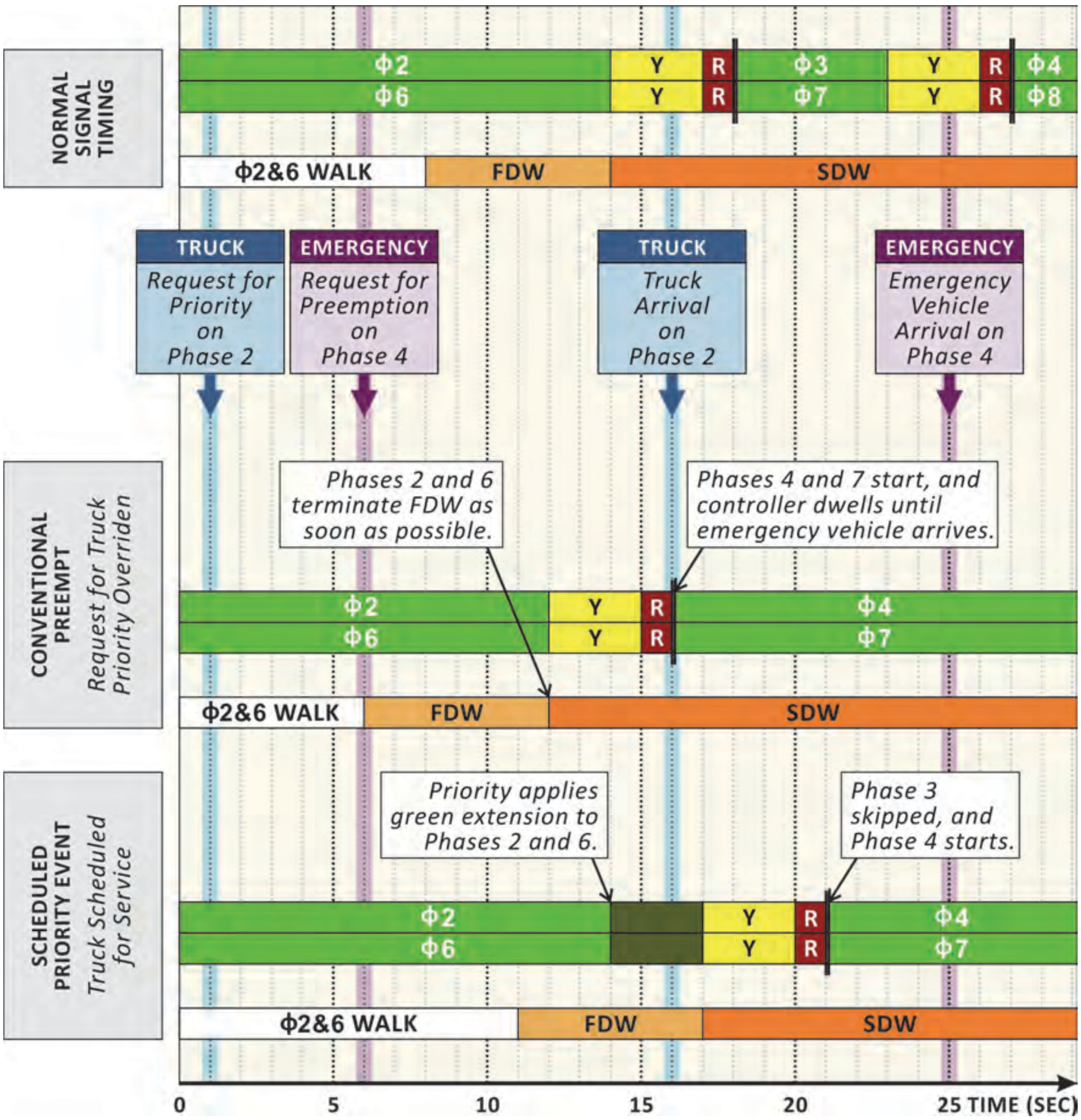
Scheduling and serving multiple preferential requests is an emerging field of practice. There are limited practical applications due to limitations in most controller firmware.

- Provision of adequate pedestrian clearance times,
- Selection of transition modes that minimize traffic disruptions, and
- Accommodation of large vehicles (e.g., trucks) on high-speed or downhill approaches.

Signal Timing Manual, Second Edition

Several researchers have evaluated the impact of scheduling requests on normal signal operations (*6*, *7*, *8*, *9*, *10*). To illustrate the schedule-based concept, Exhibit 10-11 shows how preferential treatment can accommodate a call for truck priority prior to emergency vehicle preemption. It is important to recognize that preemption is not scheduled (as it will always take precedence at an intersection), but the schedule-based approach provides more efficient transitions prior to preemption, so that other preferential treatment requests can be served.

Exhibit 10-11
Accommodating Multiple Preferential Treatment Requests through Scheduling



In this example, the service requests (from the truck and emergency vehicle) are received at different times and will require service at different times. Using schedule-based preferential treatment, the controller is able to continue serving the current phase in order to accommodate the truck, skip the next phase, and then accommodate the emergency vehicle. In order for the controller firmware to decide if both requests can be served (or if one request needs to take precedent), a TSD is required for each priority vehicle prior to its arrival at the intersection.

10.4 PREEMPTION AND PRIORITY

Every traffic signal controller implements preemption and priority in a slightly different manner. Hence, it is critical that practitioners become familiar with their firmware and shop test all applications before implementation. Section 10.4 provides information about preemption and priority controller settings that apply to all vehicle types. Specific preemption and priority controller settings that apply only to rail, emergency vehicles, transit vehicles, or trucks are discussed throughout the remainder of the chapter.

10.4.1 Preemption Settings

Preemption settings that must be selected and incorporated into the controller include but are not limited to the following (*11*):

- ***Preemption Phase(s).*** The phase (or phases) to be served as part of preemption.
- ***Limited Service (Dwell) Phases.*** All phases with movements that do not conflict with the preemption movement should be permitted to continue to operate while the signal dwells (as illustrated in Exhibit 10-12) in order to minimize delay and queuing. Limited service phases need to be identified for each preemption movement. Note that the ability to cycle through limited service phases depends on the length of the preemption event (i.e., more limited service phases will be served during longer preemption events, such as long freight trains).

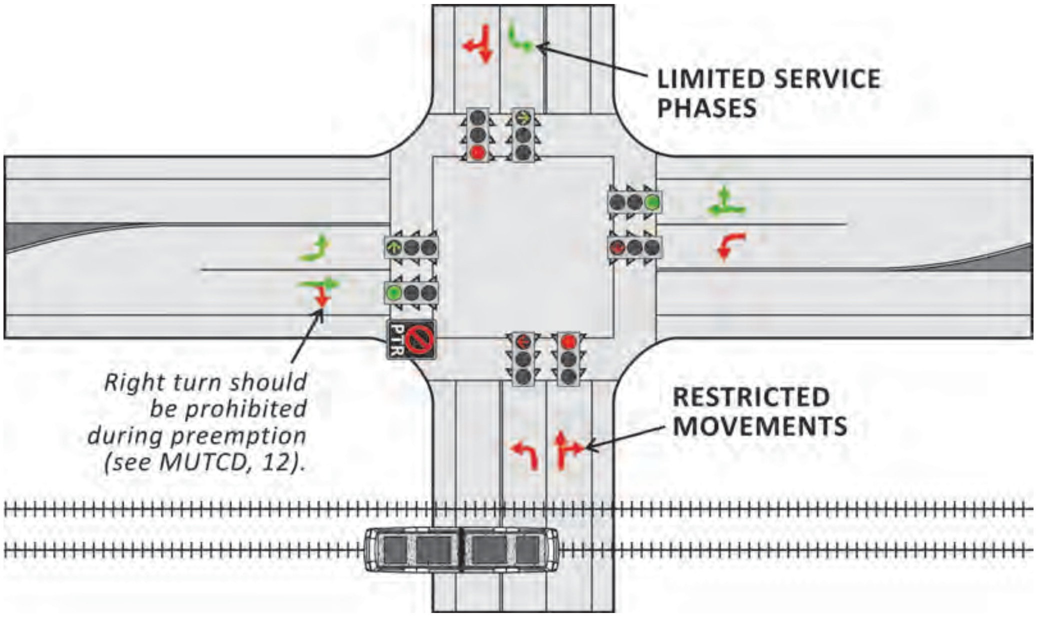
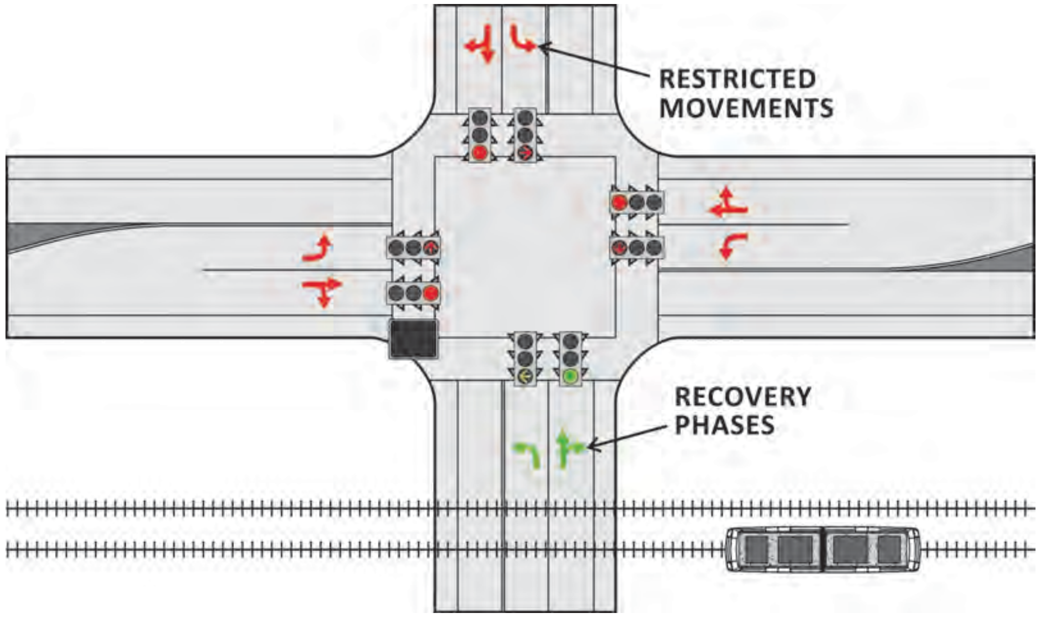


Exhibit 10-12 Limited Service (Dwell) Phases

- ***Recovery (Exit) Phases.*** Recovery phases are activated after termination of the dwell period. The phases that require rapid recovery after preemption should be designated as recovery phases (as illustrated in Exhibit 10-13).

Exhibit 10-13
Recovery (Exit) Phases



- ***Preemption Number.*** The preemption number is a unique identifier in the preemption sequence used to assign specific phases to each preemption request (e.g., rail versus emergency vehicle). Typically, six or more different preemption plans (e.g., two railroad and four emergency vehicle plans) can be implemented.
- ***Preemption Priority.*** The preemption priority is a numerical value that distinguishes the service priorities of conflicting or overlapping preemption calls on multiple approaches or for multiple users. For example, railroad preemption needs to be programmed for a higher priority than emergency vehicle preemption. Note that if more than one call of equivalent priority is received at a controller, first-come-first-serve service applies.
- ***Preemption Duration.*** Preemption duration is a set value of time for service of a preemption request. Durations may be specified by a minimum preemption duration, a maximum preemption duration, or the option to hold the preemption while the input is received (i.e., request is active for preemption). Typically, the preemption vehicle is able to check out earlier than the preemption duration.
- ***Preemption Minimum Green and Walk.*** Minimum green and walk can be set to alternate values during preemption. The minimum green during preemption should not be set less than 2 seconds.
- ***Preemption Flashing Don't Walk (FDW).*** In general, the FDW interval should be equal to the duration set under normal operations. However, the *Manual on Uniform Traffic Control Devices* (MUTCD) permits truncated FDW intervals under specific conditions (*12*).
- ***Preemption Delay.*** The preemption delay parameter delays the start of the preemption interval by a defined value. It is typically set to zero seconds. However, delay may be needed to prevent false preempts at locations where the detection is highly variable.

- ***Preemption Memory.*** Preemption memory saves the preemption request until the movement has been served. The signal should normally be operated with preemption memory active to ensure that the call is served. However, this can cause false preempts (or phantom preempts) when the detection is defective.

10.4.2 Priority Settings

When traffic signal priority is being implemented, a practitioner should consider the following priority settings (note that these may be different by mode and for low- versus high-priority requests):

- ***Priority Phasing Sequence.*** Practitioners should look at modified phasing sequences closely to ensure undesirable operations are not introduced with priority, such as a yellow trap or indications that violate user expectancy.
- ***Minimum Phase Duration.*** Practitioners should identify those phases that cannot be truncated during priority. Among the phases for which truncation is permitted, practitioners should specify the minimum duration for which a phase must be served before termination.
- ***Duration of Green/Phase Extension.*** When timing transitions are applied, practitioners should balance advantages to priority vehicles with the disadvantages to non-priority modes and movements.
- ***Minimum Green Times.*** Timing adjustments to non-priority phases should not shorten green times beyond reasonable minimum green times.
- ***Pedestrian Intervals.*** Timing adjustments to non-priority phases should not shorten pedestrian clearance (*12*).

10.5 PREEMPTION CONSIDERATIONS FOR RAIL

When a signalized intersection is located near an at-grade railroad crossing, a practitioner should consider using preemption (in conjunction with railroad warning devices) to clear any vehicular queues that may extend over the tracks (*13*), as illustrated in Exhibit 10-14. In accordance with the MUTCD, a traffic signal warrants preemption when the at-grade railroad crossing is equipped with active warning devices and is located within 200 feet of an intersection (*12*). When the at-grade crossing is farther than 200 feet from a signalized intersection, a detailed queuing analysis is recommended to determine whether signal preemption is necessary (*14*). Additional requirements are contained in the *Code of Federal Regulations* (CFR) Title 49, Section 234.225, various state laws, and the American Railway Engineering and Maintenance-of-Way Association's (AREMA's) *Communications & Signals Manual* (*15*). More information about rail preemption can also be found in the *Guide for Traffic Signal Preemption Near Railroad Grade Crossing* (*13*).

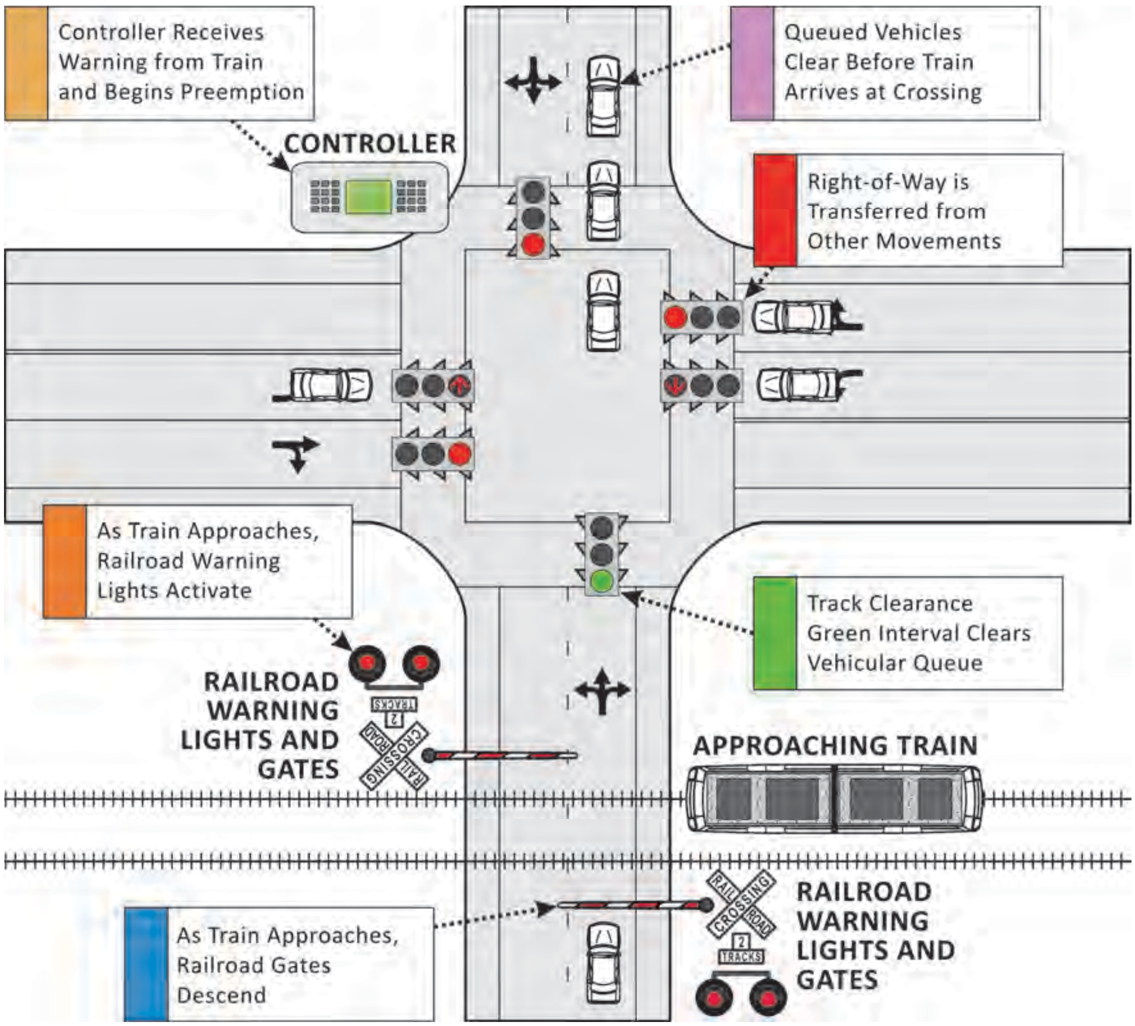
The primary purpose of rail preemption is to clear any vehicles that are stopped over the tracks before the arrival of a train.

In order for a signalized intersection to use preemption effectively to clear vehicular queues, the signal controller needs to acquire information from the railroad. Modern railroad systems use track circuits to estimate the speed and direction of a train as it enters a specified detection zone. This allows the system to predict the train's time of arrival at the at-grade crossing, which is used to calculate when the railroad warning devices (i.e., lights and gates) need to activate and when the traffic signal needs to start preemption. "Simultaneous preemption" is the most common (and most basic) type of preemption, where the track circuit activates the railroad warning devices and requests

In most cases, rail operations are kept separate from roadway operations.

traffic signal preemption at the same time. The minimum warning time required for simultaneous preemption is 20 seconds, as discussed below. However, the minimum warning time needed for the railroad warning system to activate the lights (and gates if present) may not be sufficient to clear vehicles queued over the tracks. “Advance preemption” is intended to provide additional time for a traffic signal to transition to a phase that will clear the tracks of any vehicles that might be present before the railroad warning devices are activated.

Exhibit 10-14 Railroad Preemption at an At-Grade Crossing



In order to know how much time is available for a signal to clear vehicular queues (that may be present over the tracks), a practitioner must first calculate the minimum warning time (MWT, also known as prescribed warning time) required to activate the railroad warning devices. MWT is defined as the least amount of time that warning devices shall operate prior to the arrival of a train at a railroad at-grade crossing. It is the sum of a minimum time (MT) and a clearance time (CT), as shown in Equation 10-1. According to the AREMA *Communications & Signals Manual*, the MT is a set value of no less than 20 seconds (*15*), whereas the CT is variable depending on the minimum track clearance distance. Clearance distance is measured along the highway centerline (or roadway edge if longer) from the near-side railroad warning device to a point that is 6 feet away from the far-side track. There must be 1 second of CT for every 10 feet of clearance distance greater than 35 feet. CT can also be added for site-specific conditions, such as for warning gate delay.

$$MWT = MT\ (20\ \text{seconds}) + CT\ (\text{if required})$$

Equation 10-1

where

MWT = minimum warning time (seconds),

MT = minimum time (20 seconds), and

CT = clearance time (seconds).

In the field, the observed time that the warning devices are active will often be longer than the design value of MWT because of buffer time (BT) and equipment response time (ERT) added by the railroad. BT is discretionary on the part of the railroad and may be provided in addition to MT and CT. BT is added to accommodate minor variations and ensure that the MWT is always provided. Railroad design times also account for any time required by the railroad equipment to acquire and respond to a train that enters the warning circuit. This additional time in the railroad design calculations is referred to as the ERT and may have several components depending on the complexity of the circuit (or circuits) necessary to operate the system.

*Traffic signal timing for simultaneous preemption should be based on the design value of MWT, **not** what may be observed in the field.*

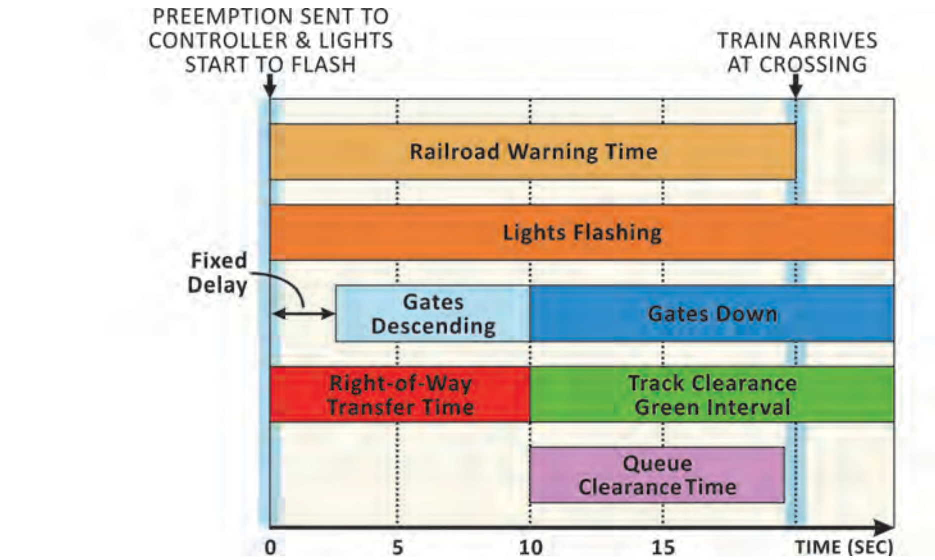
As noted previously, the most basic form of preemption is called simultaneous preemption, where notification of an approaching train is forwarded to the traffic signal controller at the same time as the design value of MWT. Modern railroad detection systems can often provide a constant warning time for this notification. In other words, railroad systems can consistently estimate when trains will arrive at at-grade crossings based on when they are detected at track circuits (as long as train speeds are relatively constant). However, areas with nearby switching are not able to provide consistent warning times (because of the variability in speeds near switching areas). These types of locations require special considerations that are beyond those discussed here and should be designed by those with detailed knowledge of the complexities.

10.5.1 Entry into Railroad Preemption

Entry into preemption is the most critical stage of railroad preemption. In this stage, the right-of-way is transferred to the "track clearance green interval" (TCGI), which is the green interval associated with the signal phase(s) that clear vehicles queued on the tracks. The time required to transfer the right-of-way to the TCGI (from the time preemption is activated in the controller) is known as the right-of-way transfer time (RTT). The RTT plus the time to clear the vehicular queue must be less than the MWT. Otherwise, vehicles could still be on the tracks when the train arrives. When the MWT for simultaneous preempt (generally limited to a maximum of 35 seconds) is inadequate to meet needs, advance preemption becomes necessary (discussed later in this section).

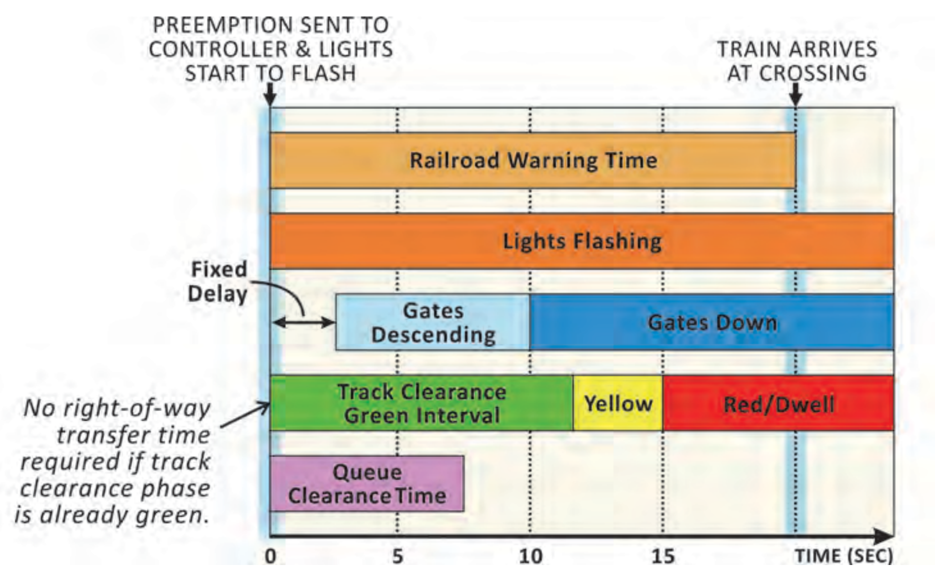
Exhibit 10-15 demonstrates the preemption entry concept as part of the simultaneous preemption process. Simultaneous preempt, while common, often requires the undesirable shortening of pedestrian clearance times. As mentioned previously, an extremely important issue is that there is a difference between calculated design values and the actual values provided by the railroad system (and thus timed in the controller). The TCGI may need to be programmed in the controller at a higher value than the calculated design value in order to account for variability.

Exhibit 10-15 Simultaneous Preemption with Maximum RTT



The components of RTT include the minimum allowable green for the current vehicle green phase, the time required for the pedestrian phase (if agency policy is to serve the pedestrian phase during preemption), and the time required for the yellow change and red clearance intervals of the active phase. If preemption is activated when the controller is in the same phase as the TCGI, the actual RTT is zero (*16*), as illustrated in Exhibit 10-16.

Exhibit 10-16 Simultaneous Preemption with No RTT



All processes described so far in this section have assumed simultaneous preemption. Simultaneous preemption is suitable when the maximum RTT plus queue clearance time can be accommodated within the designed MWT provided by the railroad. However, there are situations in which simultaneous preemption may not be adequate, such as the following:

- When it is undesirable for a pedestrian phase to be truncated.
- When longer queue clearance times require an alternative preemption strategy. (Long queue clearance times may result from substantial space between the

intersection and the crossing or when the queue has vehicles with longer start-up lost times.)

In such cases, the practitioner should request additional warning time from the railroad authorities in order to apply advance preemption (illustrated in Exhibit 10-17). The traffic signal operator may have to pay for the cost of the circuit required for the additional warning time (as a longer railroad detection zone will be needed). In the example shown in Exhibit 10-17, RTT happens to equal advance preemption time (which may or may not be the case in all situations). The example design also provides a separation time, which is desirable but not a requirement. Separation time ends the TCGI before the arrival of the train.

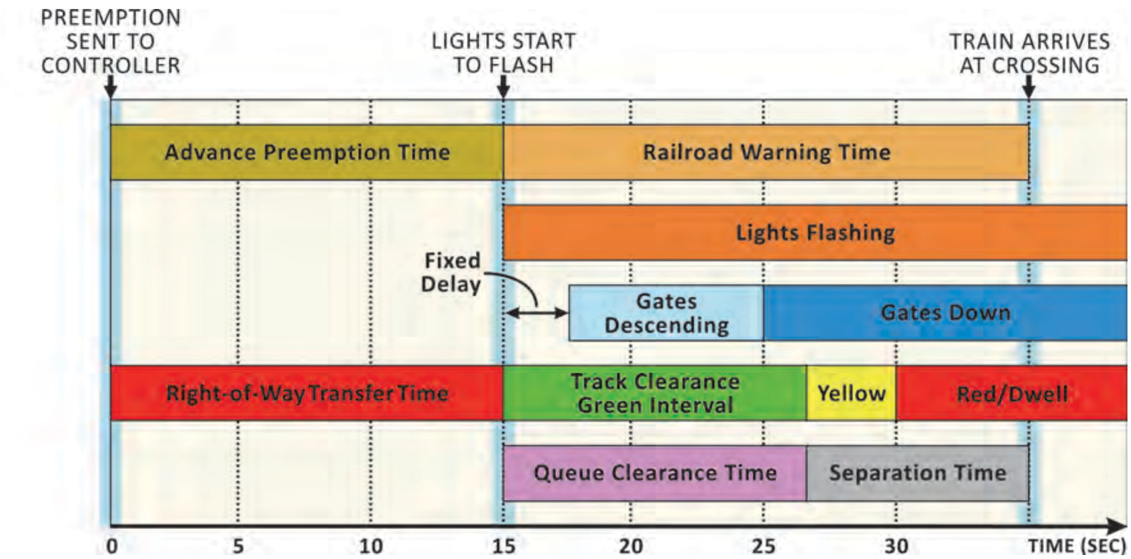


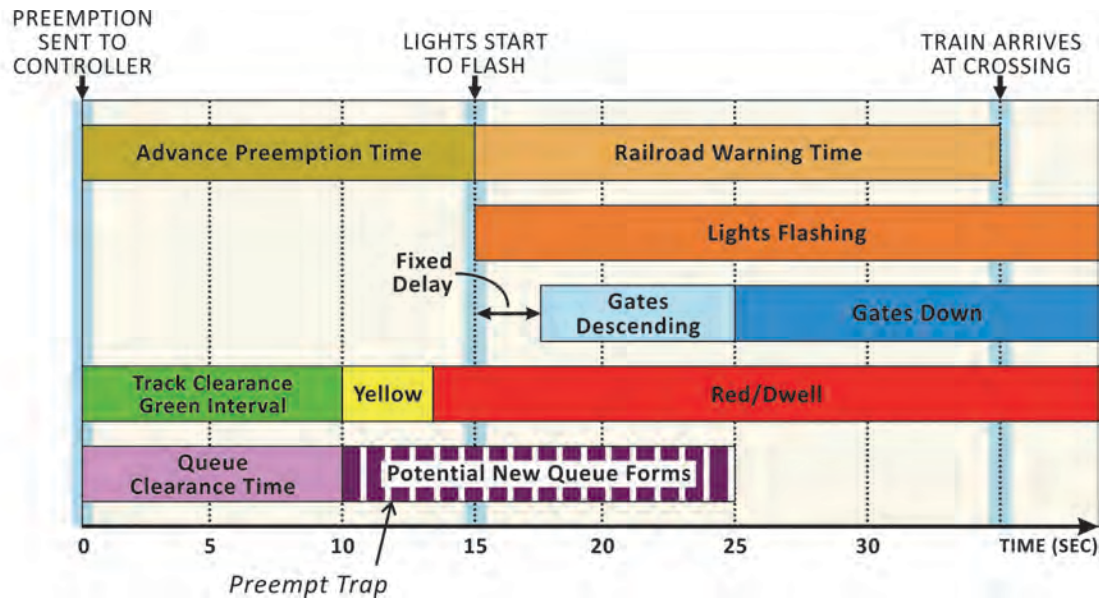
Exhibit 10-17 Advance Preemption Time

10.5.2 Advance Preemption Considerations

While an RTT of zero seconds may give a first impression of an acceptable scenario, Exhibit 10-18 illustrates that the TCGI can terminate before the lights start to flash and the gates come down (*16*). In what is referred to as the "preempt trap," a queue may form over the track before the gates descend. Once the gates lower, the queue will not be served by the controller until preemption is removed, potentially leaving the vehicles stranded on the tracks. In addition, actual advance preemption times may be longer than the design advance preemption times due to decreased speeds as the train approaches the crossing. This extended advance preemption time furthers the negative effects of the preempt trap.

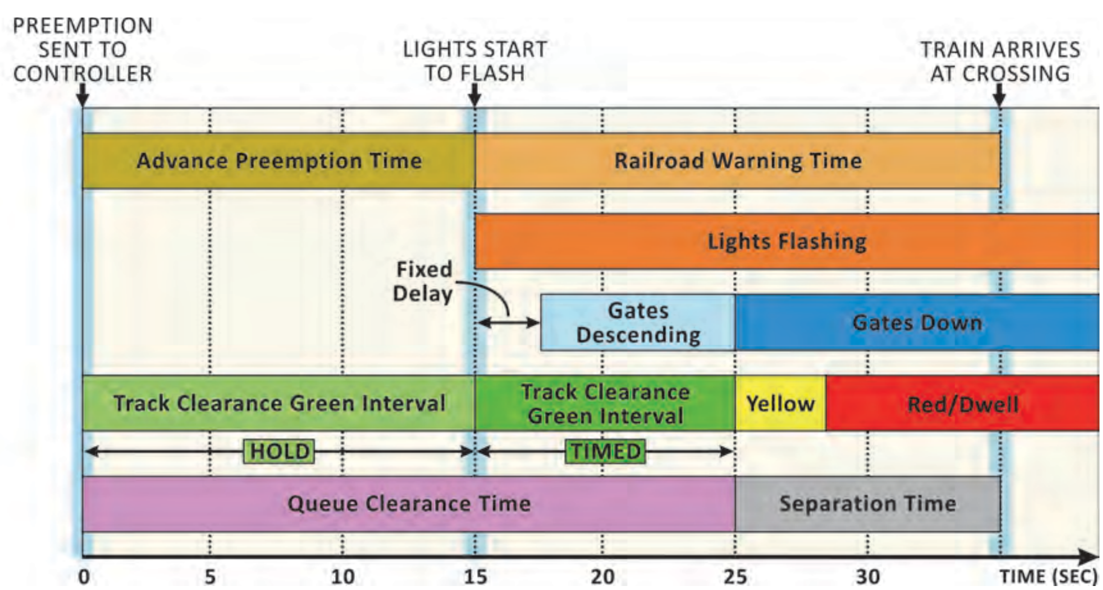
In order to eliminate the preempt trap and reduce the resultant safety concerns, several treatments can be implemented using existing railroad and traffic signal technology. One such treatment is the installation of a not-to-exceed timer by the railroad operation that forces the railroad warning devices to activate no later than the end of the design advance preemption time (*15*). Implementation of the not-to-exceed timer should be coupled with a TCGI duration that is at least equal to the advance preemption time. This strategy ensures that irrespective of both the variability in train speed and the corresponding warning time, the railroad warning devices will be activated before (or with) the TCGI, reducing the potential for vehicle queues over the tracks. In this case, a design modification is required for the railroad operations, and the traffic signal timing must be modified.

Exhibit 10-18 Example of Preempt Trap When RTT Is Zero and There Is No Mitigation



Another approach incorporates a "gate-down preempt" from the railroad signal system into the traffic signal controller (*16, 17*). With this treatment, the controller receives the conventional advance preemption call, times any RTT, and holds (dwells in) the TCGI until a second preempt is received from the railroad informing the controller that the warning gates are down. The gate-down preempt (assigned a higher preempt priority) releases the green hold resulting from the advance preempt, and enters a timed TCGI (*18*). An alternative to the gate-down preempt is the use of simultaneous preempt to time the TCGI. This alternative begins the design TCGI when the warning devices are activated. (Additional time can be added to account for the time required for the gates to descend.) Advance preemption with an RTT of zero and warning-system-active confirmation is illustrated in Exhibit 10-19.

Exhibit 10-19 Two Preempt Operation (Advance and Simultaneous)



If the controller is serving the TCGI when advance preemption is activated (as shown in the example in Exhibit 10-19), the controller dwells in the TCGI until the warning devices are activated (or gates are down) and then times the designed TCGI. This ensures that the TCGI times after the warning systems are active. This approach requires two preempts from the railroad as well as modifications to the traffic signal timing. The traffic signal controller must be capable of two preempts; the first calls and holds the TCGI, and the second implements the timed TCGI. This example is one of several designs that use more than one preempt to create better operation of traffic signal preemption.

10.5.3 Scheduling Other Calls for Service

Traditional preemption (i.e., simultaneous, advance, and advance with warning-system-active or gate-down confirmation) can have negative impacts on other phases, due to the amount of time in the TCGI. If advance preemption is considered a priority request (i.e., a request for service at a certain time scheduled in the future), then a transition into preempt can be more intelligently implemented to achieve more efficient traffic signal operations. The ability to do so depends on the functionality of the signal controller, the time required for the RTT, the amount of advance notice required, the reliability of the railroad detection system in providing consistent times, and the time remaining before critical service of the TCGI. Exhibit 10-20 illustrates the schedule-based approach, which relies on knowledge of vehicle (i.e., heavy rail or light rail transit [LRT]) arrival time, current state of signal timing service, and current detection calls.

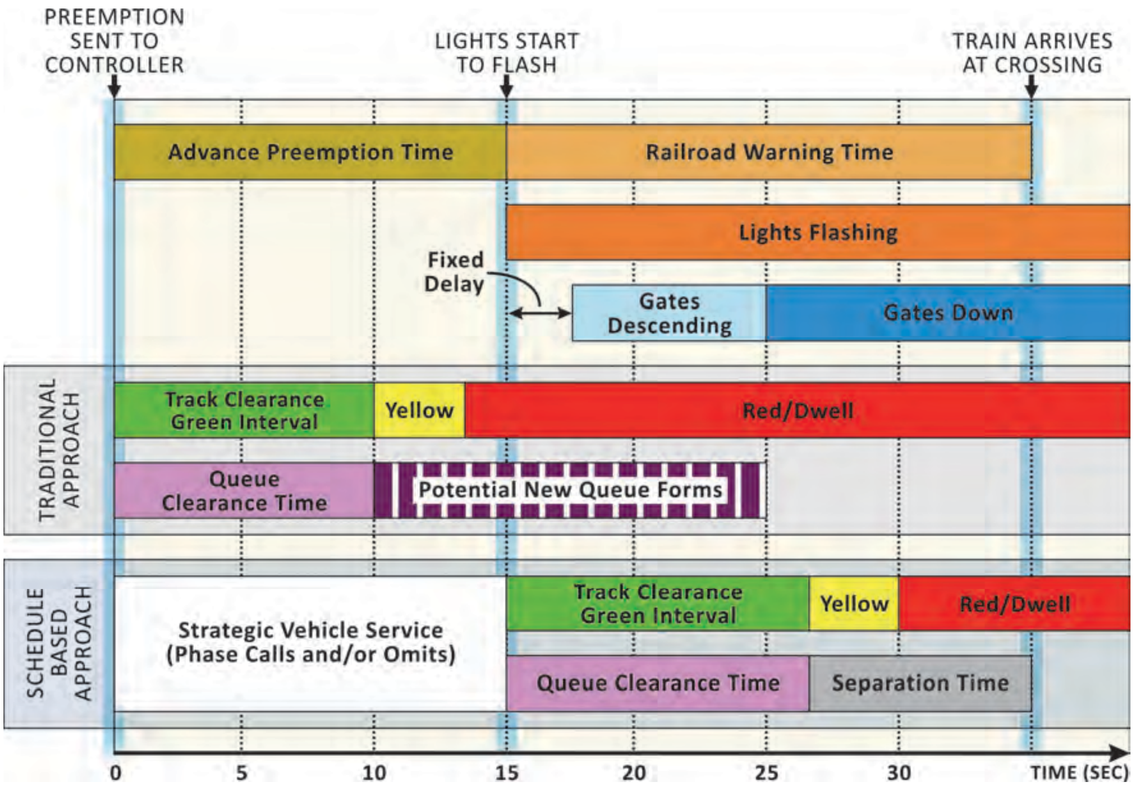
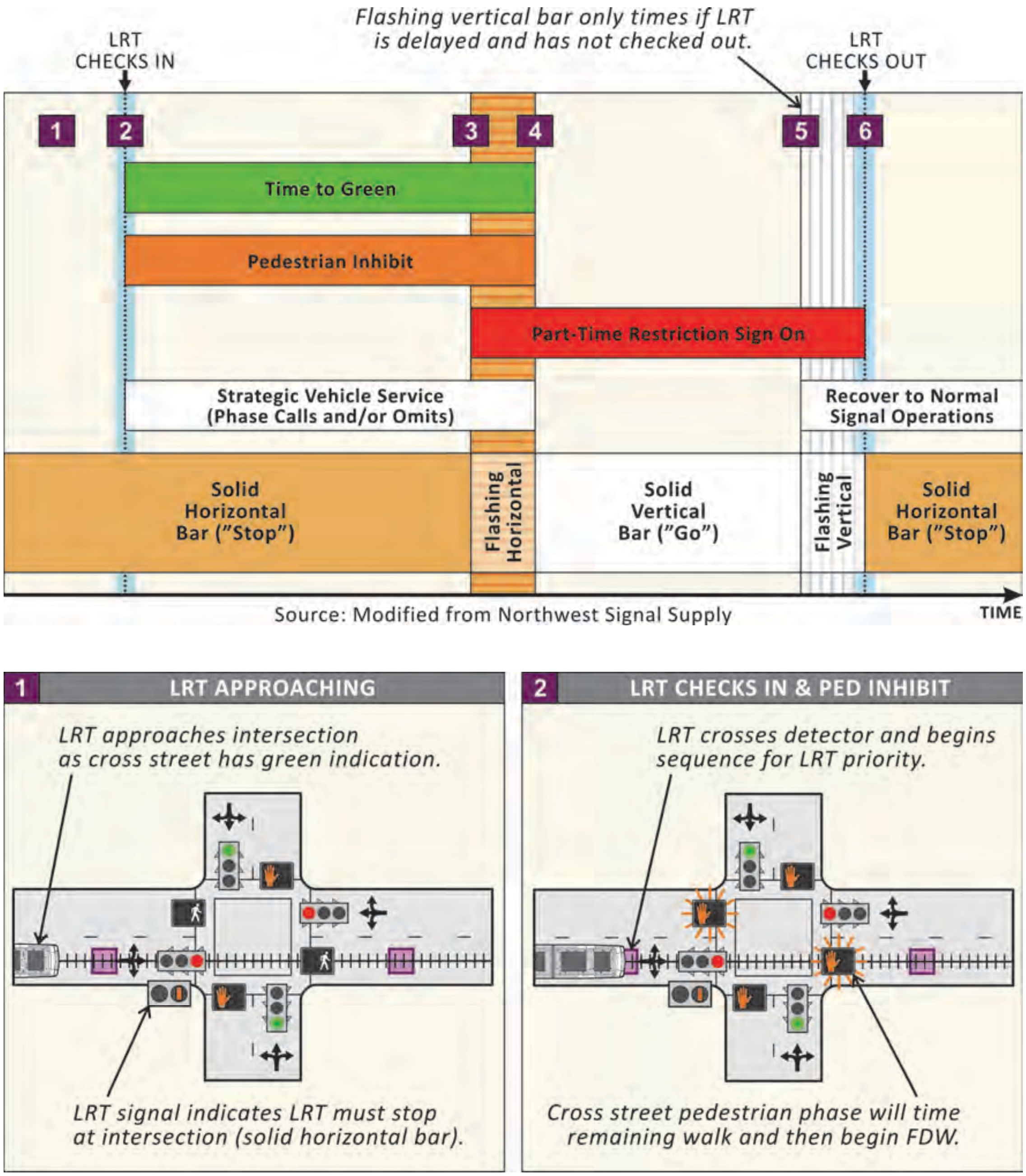


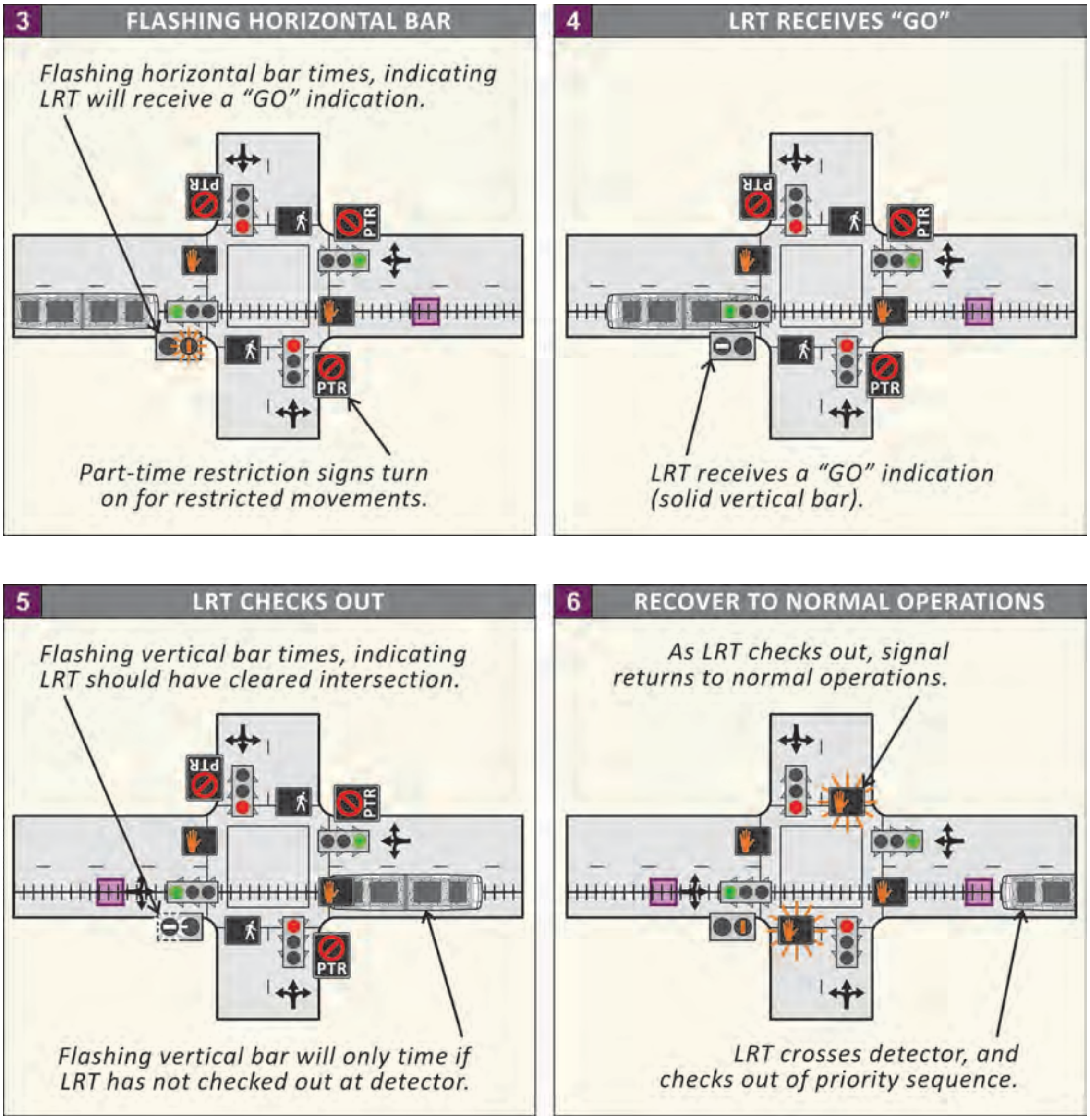
Exhibit 10-20
Schedule-Based Two Preempt Operation (Recommended Practice)

The schedule-based approach to planning for the arrival of a train has been used for LRT in Portland, Oregon, to facilitate station-to-station movements. The approach is known as “time to green” (TTG). Exhibit 10-21 (continued on the next page) illustrates this TTG scenario for LRT.

Exhibit 10-21 TTG “Scheduled” Preemption for LRT



(continued next page)



10.5.4 Railroad Preemption Dwell (or Hold)

When the TCGI times out, the controller progresses from the "entry into preemption" stage to the "dwell" (also known as "hold") stage. During the dwell stage, the signal may operate under one of the following:

- Red flash for all phases.
- Flashing yellow for major allowable movements and flashing red for allowable minor movements.
- Steady red or all-way stop.
- Limited service, in which the controller serves the allowable phases.
- Rest in green for through movements parallel to the tracks.

As mentioned previously, limited service (in which the controller serves phases not in conflict with the preemption phase) is the most preferred mode of dwell as it minimizes intersection delay.

10.5.5 Railroad Preemption Controller Settings

For the safe and efficient service of preemption requests at railroad at-grade crossings, numerous settings in the controller must be programmed. In addition to those explained in Section 10.4.1, the programmable settings necessary for the implementation of railroad preemption include, but are not limited to (*11*)

- ***TCGI.*** The TCGI is the green interval associated with the signal phase (or phases) that control movements which may queue over the tracks. A green indication (including a protected left-turn indication) should always be provided for movements associated with the TCGI.
- ***Track Clearance Green Time.*** Track clearance green time is required for the queue to clear off the tracks before the train arrives. The calculation of track clearance green time is dependent on the distance between the tracks and the intersection stop bar, the composition of the vehicles in the queue, and the relative position of those vehicles with respect to one another. The track clearance green time necessary for the safe operation of traffic signals during preemption may be calculated using a worksheet developed for the Texas Department of Transportation (*19*).
- ***Recovery (Exit) Phases.*** For railroad preemption, the controller commonly serves the phases across the tracks first in the “recovery” (exit) stage.

10.6 PREFERENTIAL TREATMENT CONSIDERATIONS FOR EMERGENCY VEHICLES

Emergency vehicle preferential treatment is used to facilitate the rapid movement of emergency vehicles through traffic signals. Often, preemption is only used for fire trucks because of the disruption to normal traffic signal operations that can occur if there are numerous service requests over short periods of time. Other emergency vehicles (e.g., ambulances and police cars) are typically not equipped with preemption capabilities. These emergency vehicles are smaller than fire trucks and more able to maneuver around vehicles in traffic. More information about emergency vehicle preemption can be found in *Traffic Signal Preemption for Emergency Vehicles: A Cross-Cutting Study* (*20*). To minimize the adverse operations effects, practitioners may implement priority routines instead of preemption.

10.7 PREFERENTIAL TREATMENT CONSIDERATIONS FOR TRANSIT

Transit signal priority (TSP) is a tool used to improve transit performance and reliability (*4*). The most common TSP strategies either extend a phase to allow a transit vehicle to pass (i.e., green extension) or terminate conflicting phases to allow early service and reduce red time (i.e., red truncation) (*6*). Green extension facilitates significantly less intersection delay than red truncation, and it should be given priority when competing calls exist. However, site-specific conditions should also be considered. For example, in order to ensure that desired objectives are met, the location of transit stops may affect how TSP is designed, timed, and implemented.

At times, signal systems with TSP may experience competing users, including multiple transit vehicles, pedestrians, bicycles, and other vehicles. Next-generation TSP can intelligently serve competing priority needs (like those shown in Exhibit 10-22) based on lateness, ridership, and route importance, as described below.

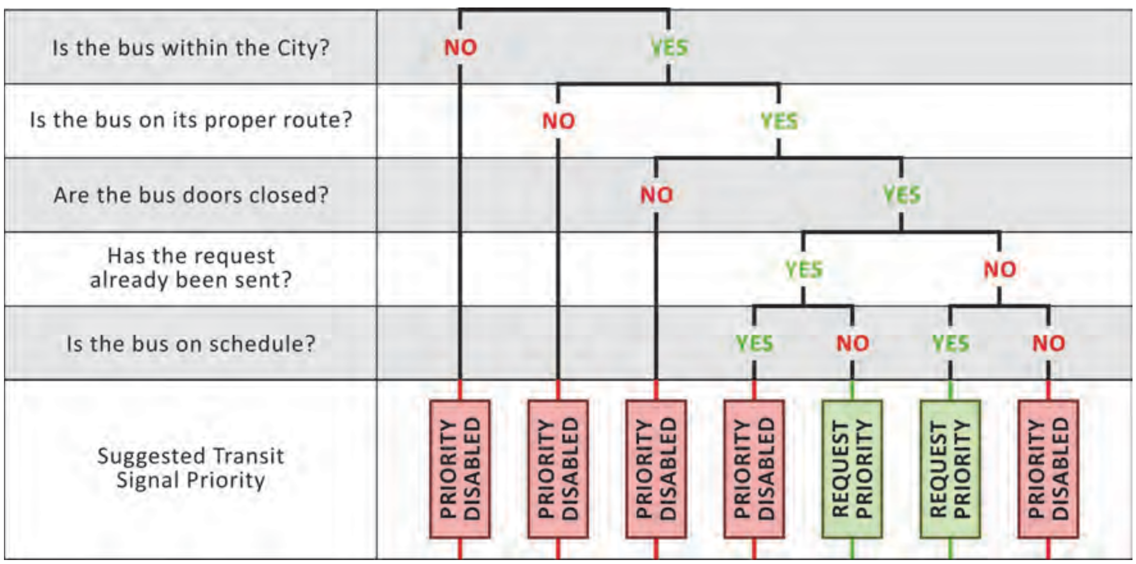
Exhibit 10-22 Multiple Transit Vehicles on Approach

- ***Schedule Adherence (i.e., Lateness).*** Practitioners often desire to provide TSP only to those transit vehicles that are behind schedule. This requires knowledge of transit vehicle locations (i.e., through use of “automatic vehicle location” [AVL] technology) and schedule information (i.e., use of computer-aided dispatch [CAD] system).
- ***Location Information.*** Practitioners occasionally desire to provide TSP only when a transit vehicle is in service, en route, and not at a transit stop. This requires high-precision GPS and knowledge of transit vehicle locations. Note that the closed position of the door on a transit vehicle is able to serve as a conditional indication for priority requests.
- ***Passenger Counter.*** Practitioners may desire to serve the route with the highest actual ridership or anticipated ridership, particularly if multiple requests for priority service occur. The ability to track historic and real-time ridership is necessary to enable this dimension of TSP.
- ***Level of Priority.*** When priority routes overlap, assigning levels of priority across vehicles and routes may be especially important to an operating agency. The system map shown in Exhibit 10-23 provides an example of a transit system for which this distinction may be necessary.

Exhibit 10-23 Example of Overlapping Transit Priority Routes

An operating agency may choose to create a decision tree that prioritizes transit vehicles based on several of the characteristics described above. Exhibit 10-24 illustrates a simple example of conditional TSP with integrated CAD/AVL data from the City of Portland, Oregon.

Exhibit 10-24 Example TSP Decision Tree



10.8 PREFERENTIAL TREATMENT CONSIDERATIONS FOR TRUCKS

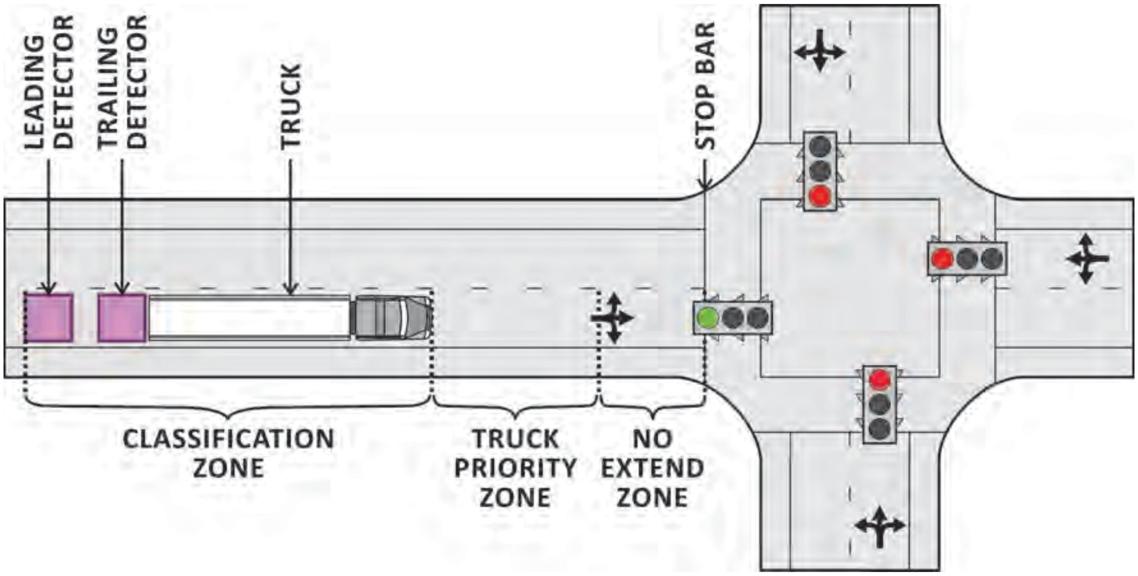
If reduced intersection delay is a desired outcome of traffic signal timing, then having a truck at the beginning of the queue is an undesirable scenario (as depicted in Exhibit 10-25). The primary objective of truck signal priority is to provide more time for trucks to pass through an intersection in order to reduce the probability that a truck will be positioned as the first vehicle at the stop bar. When used effectively, truck signal priority may result in the following benefits:

- Reduction of truck stops and delay.
- Potential reduction of truck red-light running.

- Safer phase termination for trucks (i.e., decision zone protection).
- Increased vehicular capacity of the intersection through reduced truck start-up lost time.
- Potential decrease in total intersection delay when truck priority is on a coordinated phase (as a result of the added green time to the major traffic movement).

Exhibit 10-25 Truck Position in Queue

Typically, truck signal priority is implemented on higher-speed approaches through the placement of dual upstream vehicle detectors that respond only to vehicles of a minimum length, traveling at a minimum speed. Truck drivers tend to experience longer decision zones than passenger car drivers (typically up to 8 seconds from the stop bar). For truck signal priority, effective detection is placed ***beyond***, or upstream of, the decision zone (see Exhibit 10-26). Additional information on detector considerations can be found in *Design and Installation Guidelines for Advance Warning Systems for End-of-Green Phase at High-Speed Traffic Signals* (*21*).

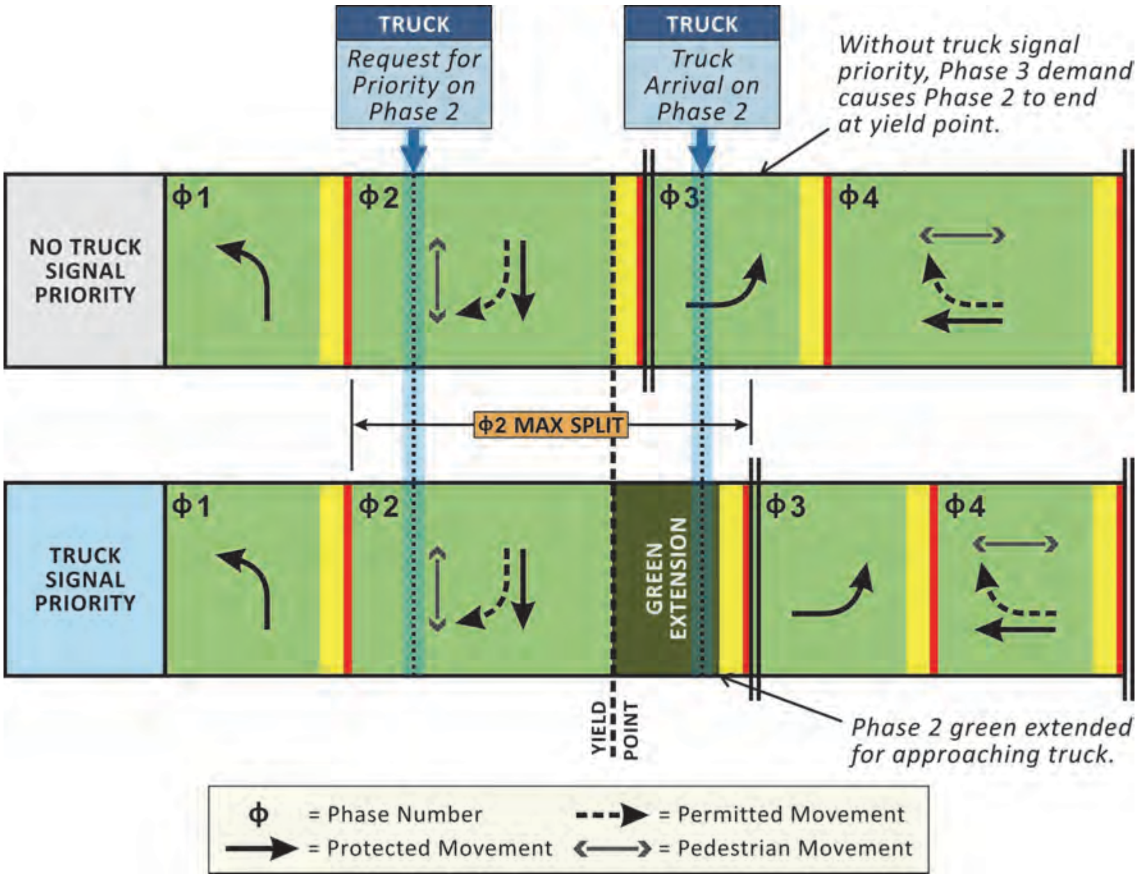


Source: Adapted from Northwest Signal Supply

Exhibit 10-26 Truck Signal Priority Detection

Truck decision zones are calculated in the truck priority algorithm, and a priority call for service is placed during that time. If the request can be accommodated, the phase is extended through a low-priority request (as illustrated in Exhibit 10-27). No early green is provided for truck signal priority, and, like TSP, only minor changes are made to signal phasing and timing if coordination is to be retained.

Exhibit 10-27 Green Extension for Truck Signal Priority



10.9 REFERENCES

1. Nelson, E. J., and D. Bullock. Impact of Emergency Vehicle Preemption on Signalized Corridor Operation: An Evaluation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1727*, Transportation Research Board of the National Academies, Washington, D.C., 2000, pp. 1–11.
2. National Transportation Communications for ITS Protocol. *Object Definitions for Actuated Traffic Signal Controller (ASC) Units*. NTCIP 1202:2005, v02.19, 2005.
3. National Transportation Communications for ITS Protocol. *Object Definitions for Signal Control and Prioritization*. NTCIP 1211, v01, 2008.
4. Smith, H. R., B. Hemily, and M. Ivanovic. *Transit Signal Priority (TSP): A Planning and Implementation Handbook*. ITS America, Washington, D.C., 2005.
5. Head, L., D. Gettman, and Z. Wei. Decision Model for Priority Control of Traffic Signals. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1978*, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 169–177.

6. He, Q., L. Head, and J. Ding. Heuristic Algorithm for Priority Traffic Signal Control. In *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2259*, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 1–7.
7. Ma, W., Y. Liu, and X. Yang. A Dynamic Programming Model for Bus Signal Priority with Multiple Requests. Paper 11-2851, Presented at 90th Annual Meeting of the Transportation Research Board, Washington, D.C., 2011.
8. Ghanim, M., F. Dion, and G. Abu-Lebdeh. Integration of Signal Control and Transit Signal Priority Optimization in Coordinated Network Using Genetic Algorithms and Artificial Neural Networks. Paper 09-3063, Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
9. Zhou, G., and A. Gan. Design of Transit Signal Priority at Signalized Intersections with Queue Jumper Lanes. *Journal of Public Transportation*, Vol. 12, No. 4, 2009, pp. 117–132.
10. Christofa, E., and A. Skabardonis. Traffic Signal Optimization with Transit Signal Priority: Application to an Isolated Intersection. Paper 11-3092, Presented at 90th Annual Meeting of the Transportation Research Board, Washington, D.C., 2011.
11. Bonneson, J., S. Sunkari, and M. Pratt. *Traffic Signal Operations Handbook,* Second Edition. Report FHWA/TX-11/0-6402-P1, Texas Department of Transportation, Austin, Texas, 2011.
12. *Manual on Uniform Traffic Control Devices for Streets and Highways,* 2009 Edition. United States Department of Transportation, Federal Highway Administration, Washington, D.C., 2009.
13. Venglar, S. P., M. S. Jacobson, S. R. Sunkari, R. J. Engelbrecht, and T. Urbanik II. *Guide for Traffic Signal Preemption Near Railroad Grade Crossing*. Report FHWA/TX-01/1439-9, Texas Department of Transportation, Austin, Texas, 2000.
14. *Preemption of Traffic Signals at or Near Railroad Grade Crossings with Active Warning Devices: A Recommended Practice*. Traffic Engineering Council Committee, TENC-4M-35. Institute of Transportation Engineers, Washington, D.C., 1997.
15. *Communications & Signals Manual of Recommended Practices,* 2012 Edition. American Railway Engineering and Maintenance-of-Way Association (AREMA), Lanham, Maryland, 2012.
16. Sun, X., T. Urbanik II, S. Skehan, and M. Ablett. Improved Highway-Railway Interface for the Preempt Trap. In *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2080*. Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 1–7.
17. Yohe, J. R., and T. Urbanik II. Advance Preempt with Gate-Down Confirmation: Solution for Preempt Trap. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2035*, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 40–49.
18. Engelbrecht, R., S. Sunkari, T. Urbanik II, S. Venglar, and K. Balke. The Preempt Trap: How to Make Sure You Do Not Have One. Report 1752-9, Texas Transportation Institute, Arlington, Texas, 2002.

19. Engelbrecht, R. J., K. N. Balke, S. R. Sunkari, and S. P. Venglar. *Research Report on Improved Traffic Signal Operation Near Railroad Grade Crossing with Active Devices*. Report FHWA/TX-03/0-4265-1, Texas Department of Transportation, Austin, Texas, 2005.
20. *Traffic Signal Preemption for Emergency Vehicles: A Cross-Cutting Study.* Report FHWA-JPO-05-010, Federal Highway Administration, United States Department of Transportation, 2006.
21. Messer, C. J., S. R. Sunkari, and H. A. Charara. *Design and Installation Guidelines for Advance Warning Systems for End-of-Green Phase at High-Speed Traffic Signals*. Report FHWA/TX-04/0-4260-2, Texas Department of Transportation, Austin, Texas, 2003.

CHAPTER 11
SPECIAL CONDITIONS

CONTENTS

LIST OF EXHIBITS

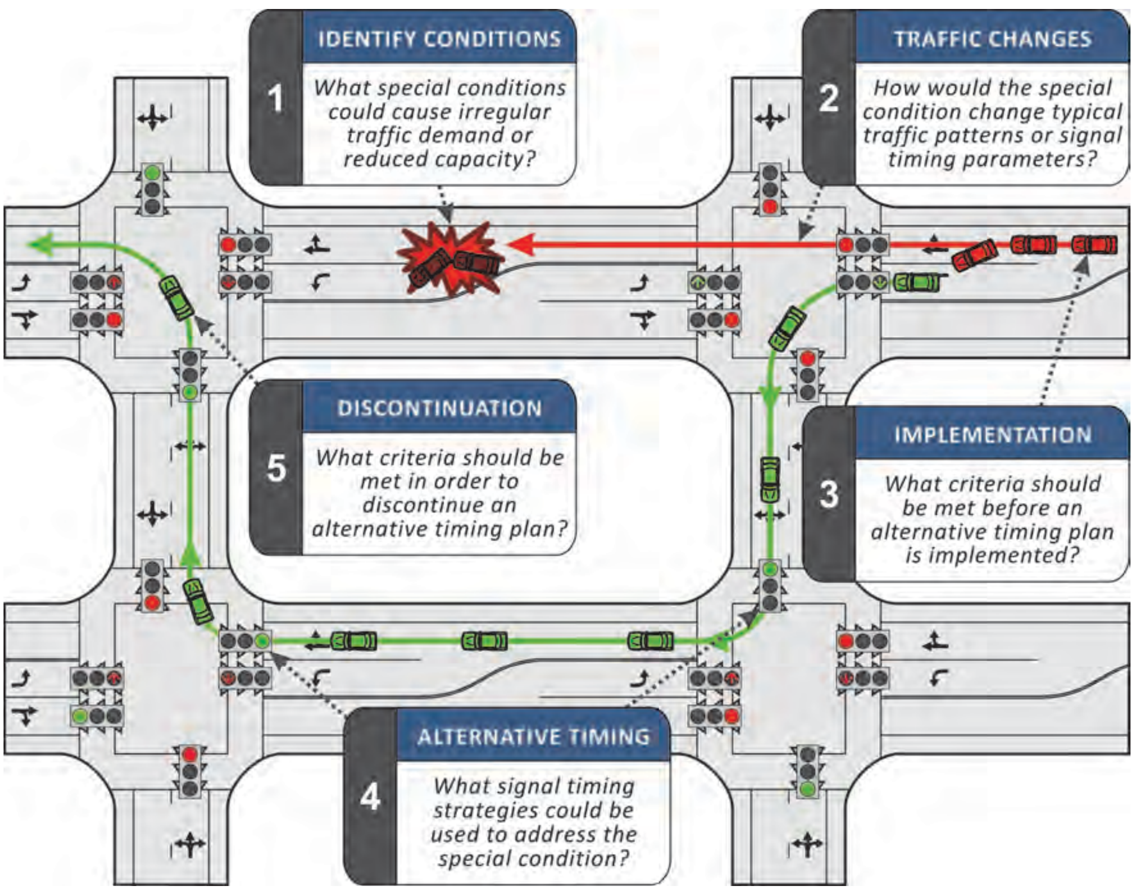
CHAPTER 11. SPECIAL CONDITIONS

The majority of signal timing plans are developed to reflect typical traffic patterns. As demand evolves over time, the plans are periodically updated. For predictable traffic patterns, this method should produce plans that effectively manage vehicle progression and queuing. For less predictable conditions, however, alternative signal timing may be required to maintain operations. In order to prepare for these special conditions, a practitioner should consider the following questions (illustrated in Exhibit 11-1):

1. What special conditions could cause irregular traffic demand or reduced capacity?
2. How would the special condition change typical traffic patterns or signal timing parameters?
3. What criteria should be met before an alternative timing plan is implemented?
4. What signal timing strategies could be used to address the special condition?
5. What criteria should be met in order to discontinue an alternative timing plan?

Additional information about each of these considerations is provided for the special conditions discussed in this chapter—weather events, traffic incidents, and planned special events.

Exhibit 11-1 Signal Timing Considerations for Special Conditions



11.1 WEATHER EVENTS

Driver behavior is directly influenced by external conditions; when weather events occur, a behavioral response is reflected in traffic operations. With the presence of fog, for example, drivers may increase following distances or reduce travel speeds to compensate for reduced visibility. In wet conditions, drivers may reduce speeds to compensate for reduced pavement friction, or they may adjust their acceleration and deceleration rates to avoid a loss of traction. Exhibit 11-2 provides a summary of potential weather-related roadway and traffic operations impacts (*1*).

Exhibit 11-2
Weather-Related Roadway and Traffic Operations Impacts

Weather Events	Potential Roadway Impacts	Potential Traffic Operations Impacts
Rain, Snow, Ice, Sleet, Hail, and Flooding	□ Reduced pavement friction □ Lane obstruction and submersion □ Reduced visibility □ Infrastructure damage □ Reduced detection capabilities □ Road restrictions and closures	□ Reduced speeds □ Increased speed variability □ Reduced roadway capacity □ Increased delay □ Increased crash risk
Strong Winds	□ Reduced visibility due to blowing snow/dust □ Lane obstruction due to wind-blown debris and drifting snow □ Reduced vehicle performance □ Reduced detection capabilities □ Bridge restrictions and closures	□ Reduced traffic speeds □ Increased delay □ Increased crash risk
Fog, Smog, and Smoke	□ Reduced visibility □ Reduced detection capabilities	□ Reduced speeds □ Increased speed variability □ Increased delay □ Increased crash risk
Lightning and Extreme Temperatures	□ Infrastructure damage	□ Traffic control device failure □ Loss of power/communications services

Agencies may develop alternative timing plans or outline specific signal timing parameter changes to manage such weather-related impacts. More specific weather-related impacts that can directly affect signal timing are summarized in Exhibit 11-3. All of these factors, and more, should be considered when adjusting traffic signal timing.

Exhibit 11-3
Weather-Related Signal Timing Impacts

Impact Category	Weather-Related Impacts
Speeds	□ Reduced vehicular travel speeds □ Reduced bicycle speeds □ Increased headways
Acceleration and Deceleration Rates	□ Reduced vehicular acceleration rates □ Reduced bicycle acceleration/deceleration rates □ Increased start-up lost times
Flow and Capacity	□ Reduced saturation flow rates □ Reduced capacity
Safety-Related	□ Increased potential for right-angle crashes
Detection	□ Reduced detection capabilities

11.1.1 Weather-Related Operations Impacts

This section expands on some of the most common operational impacts that occur as a result of weather and summarizes how they are influenced by weather severity.

Strategies that can be used to mitigate these weather-related impacts are described in Section 11.1.2.

11.1.1.1 Vehicular Travel Speeds

Freeway and arterial travel speeds can be significantly impacted by weather events, particularly events that yield precipitation (*2*). Exhibit 11-4 summarizes the effects that various weather conditions can have on vehicular travel speeds.

Exhibit 11-4
Weather-Related Reductions in Free-Flow Speeds

Roadway Surface Condition	Reduction in Free-Flow Speed (%)	
	Arterial (2)	*Freeway (3)*
Dry	0%	0%
Wet	6%	0%
Wet and Snowing	11%	13%
Wet and Slushy	18%	22%
Slushy in Wheel Path	18%	30%
Snowy and Sticking	20%	35%
Snowing and Packed	*No Data Available*	42%

11.1.1.2 Saturation Flow Rates

Studies have shown that saturation flow rates are inversely proportional to the severity of weather events. Two studies in Alaska reported significantly reduced saturation flow rates (19 percent and 12 percent reductions) during winter weather conditions (*4*, *5*). Similarly, another study with data from Minnesota reported that saturation flow rates decreased from 1800 vehicles per lane to 1600 vehicles per lane (an 11 percent reduction) during severe snowfall events with resultant accumulation (*6*).

While rain events can affect saturation flow rate, the magnitude of reduction is lower than for winter weather events. In Alabama, researchers reported that saturation flow rates dropped an average of 4.7 percent in rainy conditions (*7*). Researchers in Poland found that saturation flow rates decreased between 8.5 percent and 12.3 percent during long, all-day rainfall events; 3.6 percent during short rainfall events; 10 percent during snowfall events; and 11.4 percent with the presence of fog (*8*). Exhibit 11-5 provides a summary of reduced saturation flow rates for various weather events in Utah (*2*) and New England (*9*).

Exhibit 11-5
Weather-Related Reductions in Saturation Flow Rates

Roadway Surface Condition	Reduction in Saturation Flow Rate (%)	
	Utah (2)	*New England (9)*
Dry	0%	0%
Wet	6%	2%–3%
Wet and Snowing	11%	4%–7%
Wet and Slushy	18%	7%–15%
Slushy in Wheel Path	18%	21%
Snowy and Sticking	20%	16%

11.1.1.3 Start-Up Lost Times

While weather is often perceived to impact start-up lost time, research has found that, except for the most severe conditions, weather does not appear to have a significant effect on start-up lost time at signalized intersections. For severe weather conditions, reductions in start-up lost time were found to be most significant during wet snow conditions where accumulations on pavement surfaces occurred. The following list provides an indication of the varying degree to which wet snow conditions affect start-up lost time:

- In Burlington, Vermont, researchers studied the effects of five different weather conditions on start-up lost time. No significant differences in start-up lost time were observed during any weather conditions studied, except for when the pavement surface conditions were classified as "snowy and sticky." Under these conditions, start-up lost time increased from 2.20 seconds to 3.04 seconds (*9*).
- In Salt Lake City, Utah, only minimal increases in start-up lost time were observed during rain events (i.e., from 2.0 seconds to 2.1 seconds); however, start-up lost times increased from 2.0 seconds to 2.5 seconds when accumulations of snow were present (*2*).
- In Minnesota, start-up lost time increased from 2 seconds to 3 seconds during severe snow events. It should be noted, however, that the roadway where the study was performed was well-maintained and that traction was not an issue, as was first anticipated. On less well-maintained roadways in the area, traction was more problematic, and start-up lost times were greater (*6*).

11.1.1.4 Pedestrian Walking Speeds

Research shows that pedestrian walking speeds increase during inclement weather (*8*). In one study, the average walking speed of younger pedestrians (under 65 years) increased from 4.82 feet per second to 5.24 feet per second (a 9 percent increase). The average speed of older pedestrians (over 65 years) in the study also increased, from 4.03 feet per second to 4.37 feet per second (an 8 percent increase) (*10*). These rates are well above the design walking speed of 3.5 feet per second specified in the *Manual on Uniform Traffic Control Devices* (MUTCD, *11*).

11.1.2 Weather-Related Signal Timing Strategies

Because weather affects various aspects of traffic operations, practitioners use a number of strategies to improve safety and reduce driver error during weather events.

11.1.2.1 Increase Vehicular Red Clearance Intervals

One strategy to reduce right-angle crash potential at signalized intersections is to increase the red clearance interval between conflicting movements. Extended red clearance intervals provide additional buffer time between conflicting movements. The extension allows errant vehicles (subject to reduced pavement friction) to clear the intersection before the transfer of right-of-way to another movement. The increase in red clearance time also provides additional time for pedestrians to clear crosswalks, which may be blocked due to snow accumulation.

During weather events, the amount of additional time needed for individual clearance intervals depends on a number of factors, including the intersection geometry, approach grades, and approach speeds. Generally, no more than 1 to 2 seconds of additional red clearance time should be added for weather incidents. Practitioners should consider additional red clearance time at the following locations:

- Intersection approaches that have a documented increase in right-angle and red-light-running collisions during inclement weather.
- Intersection approaches located at the bottom of a severe downgrade.
- Minor street approaches where plowing operations are considered to be a low priority.
- Remote or isolated intersections that are difficult to reach during inclement weather.

Some controller features allow practitioners to extend the red interval dynamically until the intersection is clear. Specific logic available in some controller firmware is used to extend the red interval as long as a vehicle is located within a designated detection zone.

11.1.2.2 Increase Minimum Green Times

As previously discussed, weather can affect start-up lost time and saturation flow rates on some intersection approaches, particularly those that experience substantial snow and ice accumulation. To account for increases in start-up lost time caused by weather, some agencies increase the duration of their minimum green times on the affected approaches. Through the provision of additional green time during weather events, practitioners increase the likelihood of queue clearance under degraded conditions. This strategy is commonly deployed on intersection approaches that are situated on upgrade slopes.

11.1.2.3 Implement Phase Recalls

Weather can reduce the effectiveness of some vehicle detection systems, which can cause phases to be skipped even when demand is present. Ice cover on video detection cameras, for example, renders detectors inoperable, and fog conditions can limit the effectiveness of video detection systems. Similarly, snow accumulation can obstruct pavement markings and prompt detector error.

To mitigate this issue, some agencies place recalls on particular phases at known problem intersections prior to weather events. In the event that the detection system is rendered ineffective by weather, recalls ensure that phases continue to be served. Generally, this is done manually through remote access to the individual intersection controllers, although some video detection systems have logic processes that place constant calls to phases when fog is detected. This strategy effectively places the intersection in pretimed operations. This can limit the efficiency of operations, but will ensure that all movements are served throughout the weather event.

11.1.2.4 Weather-Responsive Coordination Plans

Many agencies activate special coordination timing plans during adverse weather conditions. The Utah Department of Transportation, for example, developed timing plans for use during considerable snow events. The timing plans were created primarily to help with snow plowing operations and to facilitate traffic flow on routes of major significance. Generally, the plans were designed to accommodate a 30 percent reduction in free-flow speed.

During snow events, on-staff meteorologists and traffic control center operators examine current weather and roadway conditions. They recommend locations and time periods for the implementation of weather-related timing plans based on the presence of one or more of the following conditions:

- A weather-related timing plan has been requested by a maintenance supervisor.
- The delay-causing portion of the weather event is expected to last more than 20 minutes.
- A significant reduction in travel speeds is detected due to weather conditions.
- A corridor is congested because of the weather event.

Because weather conditions can vary widely across a region, and even along different segments of the same roadway, weather-responsive coordination plans are implemented on a corridor-by-corridor basis. Different coordination plans may be implemented on particular corridors based on the time of day and on the intensity of the snowfall (e.g., moderate, heavy). Once a weather-related timing plan is implemented, the subject corridors are monitored via closed-circuit television by operators. When special operations are no longer needed or effective, the operators disable the timing plans.

The Connecticut Department of Transportation has also developed weather-related timing plans that are implemented with major snow events. These timing plans adjust offsets and increase cycle lengths to accommodate slower travel speeds along corridors. Control center operators implement the timing plans based on visual observation of snowfall rates and traffic conditions.

11.2 TRAFFIC INCIDENTS

Traffic incidents are unplanned, discrete occurrences, such as traffic crashes, disabled vehicles, emergency road or utility repairs, and spilled cargo. They can result in dramatic and sudden reductions to capacity (*12*), leading to excessive delay on the affected facility, and increasing the potential for secondary incidents. To effectively manage traffic incidents, an agency should develop recovery plans for mitigating the operational effects.

One of the most common traffic management strategies used during incident conditions is diverting traffic to an alternate route (*13*). Diversion can be local or regional in scale, and signal timing strategies will vary depending on the distance that vehicles are diverted. Local diversion routes are used to route traffic away from primary facilities for a short distance, typically past one point (e.g., interchange or major intersection) to the next downstream point by way of adjacent roadways. Regional traffic diversion involves strategies that deter traffic away from the bottleneck location to promote flow on multiple other facilities that serve the same downstream location (e.g., downtown area).

Alternate route plans should include information about the locations of affected traffic signals (and other traffic control devices) and the agency responsible for operating them. For locations where alternate routes extend across multiple jurisdictions, agencies may wish to develop institutional agreements or memoranda of understanding that document the conditions under which regional timing plans may be implemented and discontinued. These agreements may also define which agency is responsible for establishing, maintaining, and monitoring the incident-related timing plan.

Examples of an alternate route and associated recovery plan are provided in Exhibit 11-6 and Exhibit 11-7, respectively. Agencies must establish their own guidelines for the development of such plans, but to aid in this development, additional information about implementation, signal timing, and discontinuation strategies is provided throughout this section.

Exhibit 11-6 Example Incident-Related Alternate Route Map

Exhibit 11-7 Example Incident-Related Recovery Plan

Plan Number	10
Incident Segment	Eastbound 2nd Street
Incident Location Description	Between Street A and Street E
Affected Traffic Signals	□ Street A/2nd Street (County-operated) □ Street C/2nd Street (County-operated) □ Street E/2nd Street (County-operated) □ Street A/1st Street (City-operated) □ Street C/1st Street (City-operated)
Plan Objectives	□ Favor eastbound progression on 1st Street from Street A to Street E
Plan Assumptions	□ Assumes 1st Street is open □ Assumes access between Street A and Street E north of 2nd Street can be achieved along 3rd Street
Signal Timing Plan Changes	□ Designates eastbound right-turn at Street A/2nd Street as the coordinated phase □ Designates southbound left-turn at Street A/1st Street as the coordinated phase □ Adds 15 seconds to minimum green for southbound left-turn phase at Street A/1st Street □ Designates eastbound through at Street C/1st Street as the coordinated phase □ Designates northbound right-turn at Street E/1st Street as the coordinated phase
Implementation Criteria (Must Meet One Criteria)	□ Full closure of 2nd Street □ Peak period
Discontinuation Criteria (Must Meet One Criteria)	□ Incident removal on 2nd Street □ One-lane restoration on 2nd Street

11.2.1 Traffic Management Planning for Traffic Incidents

Agencies should define criteria and develop guidelines for determining when and where to deploy incident-related timing plans. This should be established during the planning phase to facilitate consistent decision-making and expectations among incident responders. Factors that may be used to determine when and where to implement incident-related timing plans include the following:

- Estimated duration of the incident,
- Type and severity of the incident,
- Number of lanes closed due to the incident,
- Observed traffic conditions,
- Time of day/day of week,
- Amount of available capacity along the alternate route(s),
- Traffic monitoring capabilities on the alternate route(s), and
- Type and capability of traffic signal control equipment along the alternate route(s).

Similarly, agencies should establish criteria and guidelines to determine when incident-related timing plans should be discontinued. Reasons for discontinuation of incident-related timing plans often include the following:

- Incident removal from the primary facility and full restoration of capacity.
- Partial restoration of capacity on the primary route, such that traffic demand can be accommodated on the affected facility.
- Deterioration of traffic conditions on the alternate route(s) due to secondary incidents or excessive traffic demand.

11.2.2 Incident-Related Signal Timing Strategies

Like all signal timing plans discussed in this manual, practitioners should refer to the outcome based process (introduced in Chapter 3) when developing incident-related timing plans. For many minor incidents, existing signal timing plans may be able to accommodate changes in traffic demand. This is particularly true if agencies have developed robust timing plans that can accommodate a wide range of traffic conditions. However, for major incidents, agencies may need to develop special timing plans to accommodate traffic demand on an alternate route. Common strategies used for improving traffic flow during incident conditions include the following:

Objectives should be selected for incident conditions before developing a timing strategy.

- Selecting an existing timing plan with longer cycle lengths (in order to increase the green time given to the phase that serves traffic on the alternate route).
- Implementation of a custom timing plan that features the alternate route movements.
- Deployment of a contingency “flush” plan, which consists of an extended phase or cycle to facilitate movement along the alternate route corridor.
- Using traffic management personnel or on-site technicians to manually control traffic signal operations.

After timing plan implementation, agencies should monitor the performance of the plans throughout the duration of the incident. While implementation of an incident management plan is generally more effective than no course of action, actual traffic conditions and corresponding outcomes may differ significantly from the plan. Traffic incidents, variable weather conditions, and other unexpected events may create the need for minor plan modification or for the implementation of a different signal timing plan altogether. By monitoring system performance, a practitioner can fine-tune timing parameters or implement different timing strategies as the event proceeds.

11.3 PLANNED SPECIAL EVENTS

A planned special event is defined as “a public activity, with a scheduled time and location, which impacts normal transportation system operations as a result of increased travel demand and/or reduced capacity attributed to event staging” (*14*). Examples of planned special events include sporting events, concerts, festivals, and conventions, often hosted in permanent multi-use venues (e.g., arenas, stadiums, racetracks, fairgrounds, amphitheaters, and convention centers). Major construction work zone traffic control may also warrant special traffic signal timing adjustments. Special events can include less frequent public events as well, such as parades, firework displays, seasonal festivals, and milestone celebrations at temporary venues. Potential special event impacts on various users are summarized in Exhibit 11-8 (*14*).

Exhibit 11-8 Special Event Impacts

User Class	User Type	Impact on User	Impact on Operations
Event Participant	□ Local resident □ Visitor	□ Event patron demand may cause roadway system congestion.	□ Event patrons may use another mode of travel.
Non-Attendee	□ Local resident □ Local business □ Commuter □ Truck driver □ Emergency service provider	□ Commuters and truck drivers may encounter reduced travel time reliability on corridors serving an event venue. □ Special event traffic control strategies may impact residents and businesses not involved with the event. □ Emergency service providers may experience increased response times during an event.	□ Non-attendee road users may delay planned trips or divert around a corridor impacted by a planned special event. □ Emergency service providers may mandate the provision of unimpeded emergency access routes to and from the event venue and its surrounding area.
Transit User	□ Bus rider □ Commuter rail rider	□ Transit users may experience service impacts on the day of event, including reduced availability of parking at transit stations and system capacity conditions.	□ Preferred parking areas may be set aside for commuters during the days of the event.

Planned special events differ from non-planned events (i.e., incidents) in that planned events can affect both travel demand as well as available roadway capacity. Some planned events may require road closures to accommodate vehicular and pedestrian demand. Other events, such as parades or bicycle races, may require extended street closures over a considerable distance. The extent to which planned special events affect traffic operations depends on a number of factors, including

- The type of the special event (e.g., sporting event, concert, festival, parade, race, or convention),
- The time of day and duration of the event,
- The location and size of the event venue,

- The area type in which the event occurs (i.e., urban, suburban, or rural),
- The locations from which event patrons originate,
- The scope of the event,
- The access mix (e.g., VIP, general admission) and number of patrons expected to attend the event,
- Parking availability and ease of access in the immediate vicinity of the venue, and
- Other stakeholders or influential factors (i.e., rail activity) in the vicinity of the event.

11.3.1 Traffic Management Planning for Special Events

Because event patrons are often frustrated with delays associated with event inflow and egress, event organizers prioritize the minimization of traffic congestion at planned events. In the development of traffic management strategies to accommodate special event traffic, public agencies must consider not only how patrons are impacted by the event, but also how non-attendees and transit users are affected. Event patrons generally accept a certain degree of delay as part of event attendance, but also highly prioritize arrival at the venue prior to the start of the event. Common goals when developing special event traffic management strategies include the following:

- Predictable travel to and from the event venue.
- Assurance of safety prior to, during, and after the event.
- Efficient system performance.
- Development of an emergency evacuation plan for the event area.

In order to meet these goals, traffic management plans commonly include the following components (*14)*:

- Site access and parking plan,
- Pedestrian access plan,
- Traffic flow plan,
- Traffic control plan,
- En-route traveler information plan,
- Traffic surveillance plan, and
- Traffic incident management and safety plan.

The traffic management plan should be debriefed after the event ends. The purpose is not only to assess areas for improvement, but also to identify successes and efficiencies in the overall traffic management plan. Post-event debriefs should be arranged during the event planning. Agencies that do not plan a post-event debrief may find it difficult to obtain full participation by all stakeholders. The meeting should occur within a few days of the event. This delay is long enough that event planners and stakeholders can resolve personal assessments of the event, but short enough to reduce memory recall failures.

11.3.2 Special-Event-Related Signal Timing Strategies

Event organizers frequently request that law enforcement manually control traffic signals on the day of an event, but versatile traffic signal timing plans can often perform just as well (if not better) than traffic control officers (except in extreme cases). While the manual operation of traffic signals can assist patrons with access to and from the event venue, it can also disrupt planned flow patterns and coordination at adjacent signals.

Special-event-related plans may prioritize either major or minor street traffic movements. Similar to traffic incident management, common timing plan strategies for planned special events include the following:

- Use of an existing timing plan with longer cycle lengths to increase the length of normally favored phases.
- Implementation of custom timing plans that favor movements to and from the venue.
- Deployment of a contingency “flush” plan that includes an extended phase or cycle to facilitate movement through a corridor.
- Manual traffic signal system operator control to increase time for a movement.

Increasing roadway capacity can be done in combination with these traffic signal timing strategies to further mitigate operations. Tactics for increasing roadway capacity during a special event include the following:

- Restriction of on-street parking near the event venue.
- Utilization of shoulders as vehicle travel lanes.
- Use of alternative lane configurations/operations, including reversible lane operations and contraflow operations.
- Restriction of commercial street access to business employees, customers, emergency vehicles, taxis, and transit buses.
- Alternate route deployment for background through traffic and event-generated traffic to facilities around the restricted street.

Agencies should monitor the performance of the plans throughout the duration of the special event. To manage traffic effectively, practitioners need timely, accurate information. Among other objectives, real-time monitoring of traffic conditions allows signal operators to

- Track changes in system performance during the event.
- Identify locations or corridors with poor performance.
- Identify potential causes and associated remedies (i.e., contingency plans).
- Identify specific areas that require improvement/enhancements for future events.
- Provide information to decision-makers and the public.
- Provide input to post-event evaluation.
- Assist emergency responders as necessary.

For major special events, event planners and agencies will often use traffic management teams to adjust traffic management plans in real-time and to launch contingency scenarios as the events unfold. A minimum of one traffic signal technician on the traffic management team is recommended on the day of the event to perform emergency maintenance and to make on-site adjustments to deployed strategies. Traffic signal technicians should also be available to facilitate the timing plan downloads as conditions change, while mobile technicians can make quick, on-site changes at critical locations to meet user needs on the day of the event.

11.4 REFERENCES

1. Park, B., T. K. Jones, and S. O. Griffin. *Traffic Analysis Toolbox Volume XI: Weather and Traffic Analysis, Modeling and Simulation.* Report FHWA-JPO-11-019, Federal Highway Administration, United States Department of Transportation, 2010.
2. Perrin, H. J., P. T. Martin, and B. G. Hansen. Modifying Signal Timing During Inclement Weather. In *Transportation Research Record: Journal of the Transportation Research Board*, *No. 1748*, Transportation Research Board, National Research Council, Washington, D.C., 2001, pp. 66–71.
3. McBride, J. C., W. J. Kennedy, J. H. Thuet, M. C. Belangie, R. M. Stewart, C. C. Sy, and F. R. McConkie. *Economic Impact of Highway Snow and Ice Control.* Report FHWA-RD-77-95, Federal Highway Administration, United States Department of Transportation, 1977.
4. *Anchorage Signal System Upgrade—Final Report.* Bernardin Lochmueller & Associates, Inc., Anchorage, Alaska, 1995.
5. Botha, J., and T. Kruse. Flow Rates at Signalized Intersections under Cold Winter Conditions. *Journal of Transportation Engineering,* Vol. 118, No. 3, 1992, pp. 439–450.
6. Maki, P. J. Adverse Weather Traffic Signal Timing. *Proc., 69th Annual Meeting of the Institute of Transportation Engineers*, Las Vegas, Nevada, 1999.
7. Sullivan, A. J., V. P. Sisiopiku, and S. Reddy-Surhabi. *Impacts of Signal System Timings on Rain Related Congestion.* Report FHWA/CA/OR, University Transportation Center for Alabama, 2010.
8. Chodur, J., K. Ostrowski, and M. Tracz. Impact of Saturation Flow Changes on Performance of Traffic Lanes at Signalised Intersections. Presented at the 6th International Symposium on Highway Capacity, Stockholm, Sweden, 2011.
9. Agbolosu-Amison, S. J., A. W. Sadek, and W. ElDessouki. Inclement Weather and Traffic Flow at Signalized Intersections: Case Study from Northern New England. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1867*. Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 163–171.
10. Knoblauch, R., M. Pietrucha, and M. Nitzburg. Field Studies of Pedestrian Walking Speed and Start-Up Time. In *Transportation Research Record 1538*. Transportation Research Board, National Research Council, Washington, D.C., 1996, pp. 27–38.

11. *Manual on Uniform Traffic Control Devices for Streets and Highways,* 2009 Edition. United States Department of Transportation, Federal Highway Administration, Washington, D.C., 2009.

12. Owens, N., A. Armstrong, P. Sullivan, C. Mitchell, D. Newton, R. Brewster, and T. Trego. *Traffic Incident Management Handbook.* Report FHWA-HOP-10-013, Federal Highway Administration, United States Department of Transportation, 2010.

13. *Alternate Route Handbook*. Report FHWA-HOP-06-092, Federal Highway Administration, United States Department of Transportation, 2006.

14. Latoski, S. P., W. M. Dunn, Jr., B. Wagenblast, J. Randall, and M. D. Walker. *Managing Travel for Planned Special Events.* Report FHWA-OP-04-010, Federal Highway Administration, United States Department of Transportation, 2003.

CHAPTER 12
OVERSATURATED CONDITIONS

CONTENTS

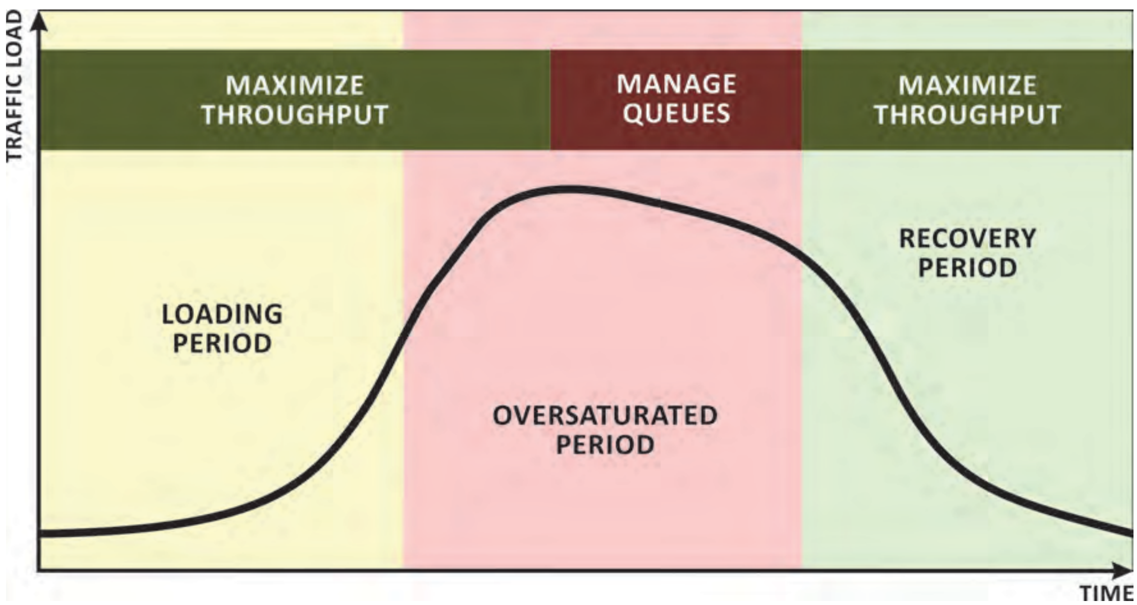
LIST OF EXHIBITS

CHAPTER 12. OVERSATURATED CONDITIONS

Oversaturated conditions occur when demand exceeds capacity, which is frequently observable with heavy congestion and the presence of long queues. Under such conditions, operational objectives are commonly maximizing throughput and managing queues (*1*), but practitioners may find it difficult to develop effective traffic signal timing to accommodate those operational strategies. Although the issue of oversaturation is often not resolvable solely through new signal timing, there are several mitigation strategies that can be applied to improve overall system performance and increase short-term capacity.

When oversaturated conditions are present, performance is typically evaluated during three periods: the loading period, the oversaturated period, and the recovery period. These performance periods (see Exhibit 12-1) are assigned based on traffic load and, depending on the operating environment, may last for several minutes to several hours. When oversaturated conditions persist for an extended period of time, it may be beneficial to apply different timing strategies during the three performance periods.

Exhibit 12-1
Oversaturation Performance Periods



Characterizing the situation is the first step when selecting a mitigation strategy. This requires the practitioner to define the type of oversaturation, using the following characteristics:

- Number of affected intersections,
- Number of affected directions of travel,
- Duration of oversaturation,
- Degree of change over time,
- Frequency of oversaturation,
- Causes of oversaturation, and
- Specific symptoms of oversaturation.

The following sections will help practitioners define the type of oversaturation that is present, the potential operational objectives specific to oversaturation, and the operational strategies for accomplishing those objectives.

12.1 SYMPTOMS OF OVERSATURATION

Exhibit 12-2 provides a summary of many of the common symptoms that indicate oversaturated conditions. One of the reasons that oversaturation is such a challenging problem for traffic signal control is that these symptoms often occur in combination. Identifying symptoms (and the degree to which they are present) is important when selecting appropriate operational objectives and mitigation strategies. Additional information about each oversaturation symptom is provided throughout this section.

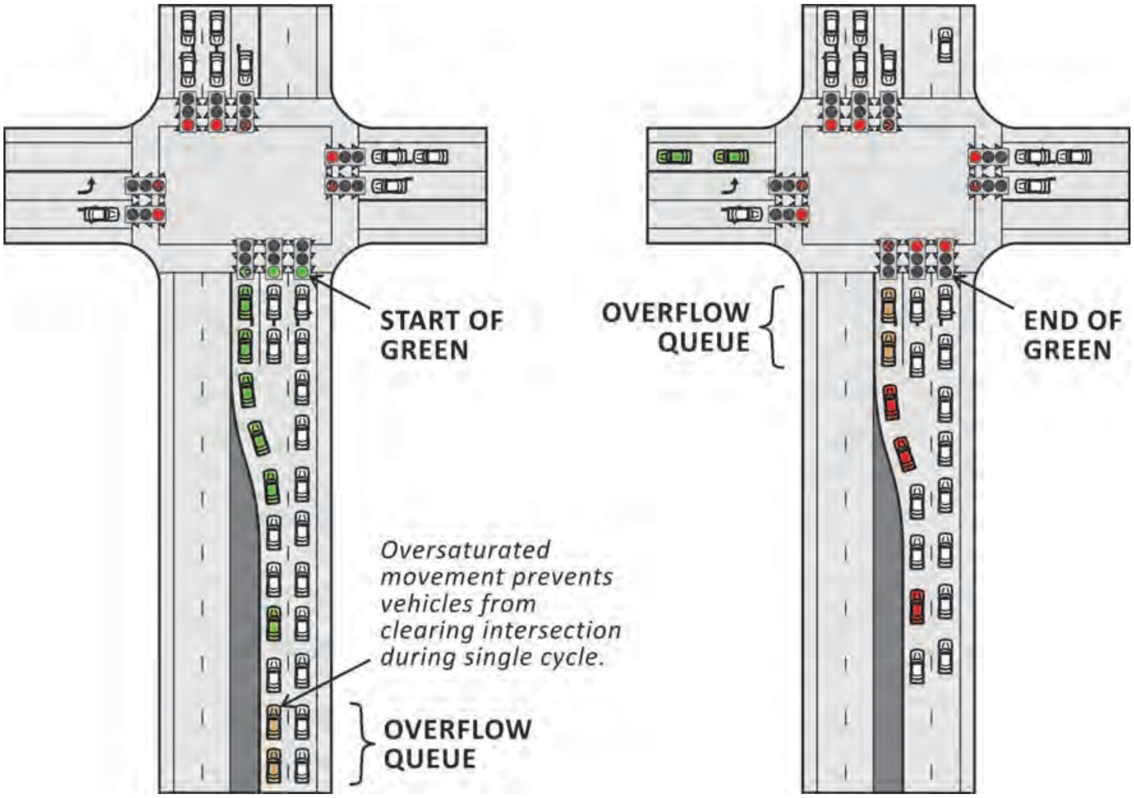
Exhibit 12-2 Oversaturation Symptoms

Oversaturation Symptom	Description
Overflow Queue	Part of queue not served during a single cycle
Approach Spillback	Upstream queue physically blocks downstream vehicles
Storage Bay Spillback	Turning queue fills storage bay and physically blocks through movement
Storage Bay Blocking	Through queue physically blocks turning movement from storage bay
Starvation	Traffic demand is restricted from using full downstream roadway capacity
Cross Intersection Blocking	Queue extends into an intersection and blocks progression of crossing vehicles

12.1.1 Overflow Queue

An overflow queue is one of the most easily observed indications of oversaturated conditions. It is defined as the part of a queue that is not processed during the green interval and that must be served by subsequent cycles (as shown in Exhibit 12-3).

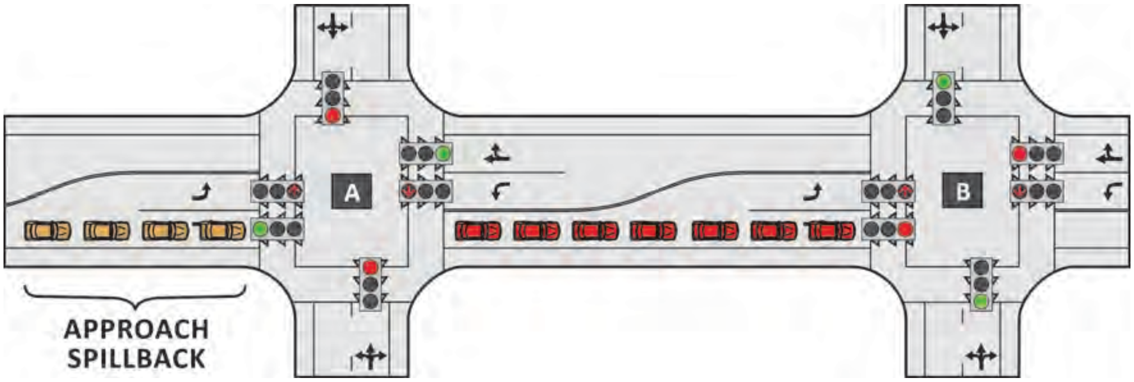
Exhibit 12-3 Overflow Queue



12.1.2 Approach Spillback

Approach spillback occurs when the queue from a downstream intersection occupies all available space on a link and prevents upstream vehicles from entering the downstream link on green (see Exhibit 12-4). Some literature has defined this condition as "de facto red" for the upstream movement because no progression is possible. The oversaturation cause is Intersection B, and simple timing adjustments at Intersection A (without understanding the problem at Intersection B) are likely to be ineffective.

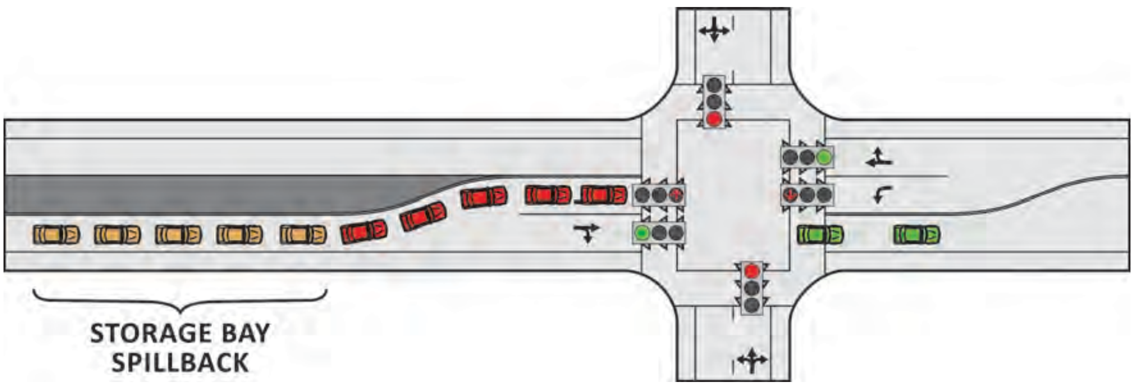
Exhibit 12-4 Approach Spillback



12.1.3 Storage Bay Spillback

Storage bay spillback (shown in Exhibit 12-5) occurs when turning traffic uses up the entire space of the storage bay and blocks through traffic. The blocked through movement may then experience starvation (explained in Section 12.1.5). Storage bay spillback is commonly caused by inadequate green time for the turning phase or ineffective phase sequencing.

Exhibit 12-5 Storage Bay Spillback



12.1.4 Storage Bay Blocking

Storage bay blocking (illustrated in Exhibit 12-6) is the converse of storage bay spillback. Storage bay blocking occurs when queues for a through movement extend beyond the opening of a storage bay. In this situation, the turning movement experiences starvation (explained in Section 12.1.5) because left- or right-turning vehicles are blocked. If the storage bay is unoccupied because through vehicles have blocked all turning traffic, the turning movement may be skipped entirely. This condition is commonly caused by inadequate green time for the through movement or by ineffective phase sequencing.

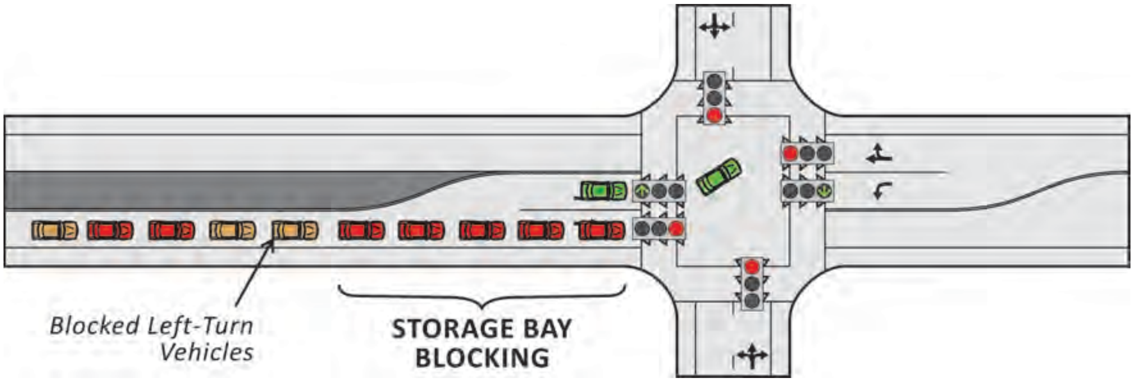


Exhibit 12-6 Storage Bay Blocking

12.1.5 Starvation

Starvation occurs during the green interval when the use of full roadway capacity is restricted by the effects of spillback, storage bay blocking, or perhaps because the upstream signal is red. In this situation, one of these limiting factors starves the downstream roadway of traffic demand (as illustrated in Exhibit 12-7). This is commonly a result of improper offsets or phase sequence.

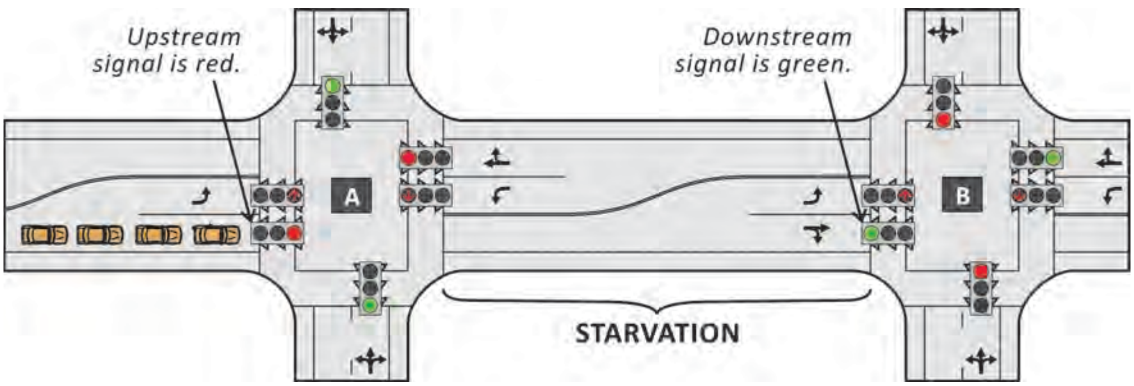


Exhibit 12-7 Starvation

12.1.6 Cross Intersection Blocking

Cross intersection blocking (illustrated in Exhibit 12-8) occurs when queues extend into an intersection and block the progression of crossing vehicles. While most jurisdictions have "don't block the box" laws or policies, these situations are not uncommon in grids and networks with short link lengths. Carefully controlled green time settings and signal offsets are necessary when mitigating this scenario.

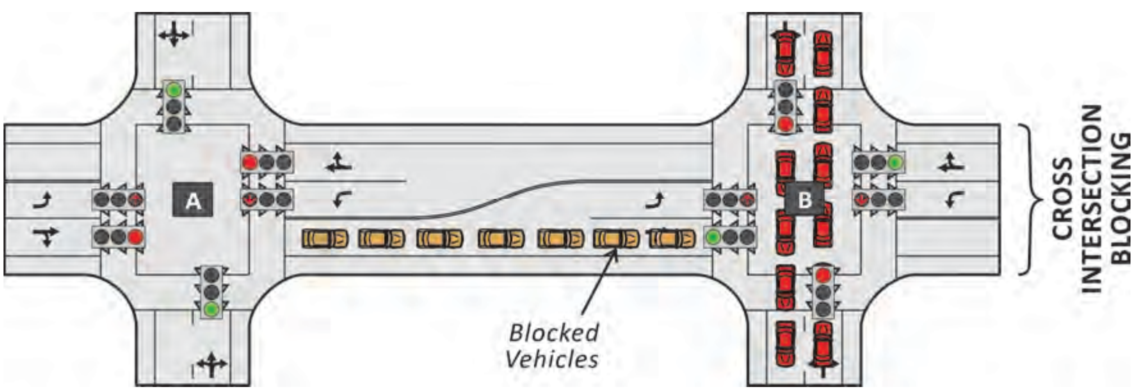


Exhibit 12-8 Cross Intersection Blocking

12.2 OPERATIONAL OBJECTIVES FOR OVERSATURATED CONDITIONS

When conditions are oversaturated, the operational objective is usually to maximize throughput or to manage queues. Minimizing user delay is not typically a viable objective in oversaturated conditions, as it is no longer possible to avoid phase failures. Maximizing throughput and minimizing queues, however, keep as much of the system operational as possible.

12.2.1 Maximizing Intersection Throughput

Intersection throughput is defined as the number of vehicles that can be served by an intersection over a specified period of time when continuous demand exists. It is typically assessed using two measurements:

- The total number of vehicles “input” to a system of intersections and
- The total number of vehicles “output” by a system of intersections.

When running efficiently, a signal network will experience an input rate and output rate that are nearly equal. In oversaturated conditions, however, the output rate is less than the input rate, and overflow queues begin to build at various points within the system.

Maximizing throughput manages the formation and growth of queues, which prevents spillback and resultant congestion. Situations do arise in which pervasive queues will continue to grow until demand diminishes, and further revisions to signal timing are futile. However, the practitioner should consider the following strategies (described in detail throughout Section 12.3) during the performance periods:

- Loading Period—Maximizing both ***input*** and ***output***.
- Oversaturated Period—Using available physical capacity to maximize ***input***.
- Recovery Period—Dissipating queues to maximize ***output***.

12.2.2 Queue Management

When maximizing throughput fails to relieve growing residual queues, many practitioners change their objective to managing queues. Effective queue management is accomplished through the use of signal timing strategies that allow queues to form in locations where they are least likely to have a detrimental effect on the overall operation of the corridor or network. Specific strategies for queue management are outlined in Section 12.3.

12.3 MITIGATION STRATEGIES FOR OVERSATURATED CONDITIONS

This section describes intersection, arterial, and network mitigation strategies that can be used to relieve oversaturated conditions. Exhibit 12-9 provides a summary of the various mitigation strategies and whether their primary objective is to maximize intersection throughput or manage queues. Depending on the system and oversaturated conditions, these strategies may be used in combination or sequence.

Exhibit 12-9
Oversaturation Mitigation Strategies

Operational Objective	Type of System	Mitigation Strategy	Definition
Maximize Intersection Throughput	Individual Intersection(s)	Split Reallocation	Reallocating split time from under-saturated phases to oversaturated phases.
		Cycle Length Increase	Adding split time to oversaturated phases without reducing split time for minor movements, effectively increasing the cycle length.
		Operation of Closely Spaced Intersections on One Controller	Operating two intersections using one controller, allowing close coordination.
	Arterial	Phase Sequence Modification for Left Turns[1]	Changing the phase sequence to lead or lag left turns, increasing the bandwidth for progressing platoons.
		Preemption Flushing (“Green Flush”)	Progressing oversaturated movements by placing calls from downstream intersections to upstream intersections.
		Alternative Timing Plan Flushing	Using a timing plan to increase the cycle length and give the majority of the green time to the oversaturated phases.
Manage Queues	Individual Intersection(s)	Green Extension	Adding green time to a phase when a detector exceeds a defined threshold of occupancy.
		Phase Re-Service	Serving an oversaturated phase twice during the same cycle.
		Phase Truncation	Termination of a green interval (despite demand) when there is minimal to no flow over a detector.
	Arterial	Simultaneous Offsets	Setting offset to zero between intersections, so that queues move simultaneously.
		Negative Offsets	Starting the green interval earlier at downstream intersections to allow the downstream queue to dissipate before upstream vehicles arrive.
		Offsets to Prevent Queue Spillback	Choosing offsets that allow the downstream queue to clear before the upstream vehicles reach the end of the downstream queue.
		Offsets to Prevent Starvation	Choosing offsets that ensure the first released vehicle at the upstream intersection joins the discharging queue as it begins to move.
	Network	Metering (Gating)	Impeding traffic at appropriate upstream points to prevent traffic flow from reaching critical levels at downstream intersections.
Maximize Intersection Throughput and Manage Queues	All	Adaptive Control	Applying detection data and adaptive signal control algorithms to adjust signal timing.
		Combination of Mitigation Strategies	Combining mitigation strategies and/or using them in sequence throughout the oversaturated period.

[1] This strategy can also be applied at the individual intersection level, but is primarily used for arterials.

12.3.1 Mitigation Strategies for Individual Intersections

Some strategies are best utilized at individual intersections when mitigating oversaturated conditions.

12.3.1.1 Split Reallocation

When oversaturated conditions exist, it is often necessary to reallocate splits to provide a larger portion of the cycle to oversaturated phases and shorter splits to under-saturated phases. While reallocating split time to accommodate demand on a saturated movement is detrimental to minor movements, the overall benefit is usually recovered through a reduction in delay on the oversaturated approach. Even just a few seconds of reallocated green time can noticeably reduce the growth of overflow queues.

By observing the arrival rate and the departure capacity of the oversaturated movement, practitioners can typically calculate the amount of green time required to reduce overflow queues. Phases with lower volume-to-capacity (v/c) ratios are best suited for reallocation of split time, but care should be taken to keep the volume-to-capacity ratios below 1.0. If this is not possible, the phase (or phases) with the most upstream storage space for queued vehicles should be selected for split-time reduction.

12.3.1.2 Cycle Length Increase

When split reallocation causes an excessive dis-benefit to minor movements, an increase in cycle length may be required in order to obtain the green time needed for oversaturated movements. This increase does not mean that the cycle length must be long. A common misconception held by many practitioners is that long cycle lengths are needed during oversaturated conditions, which leads to the tendency to increase cycle lengths as queue lengths increase. This misconception is founded on the notion that longer cycle lengths are more efficient because a smaller portion of the cycle is lost due to phase changes. However, research and practical experience have shown that when green intervals are longer than 30 seconds, saturation flow rates and overall intersection efficiency decline (*1*). Therefore, when all approaches at an intersection are oversaturated, the most efficient cycle length is one that terminates service of a phase when the saturation flow rate drops, so that subsequent oversaturated phases may be served.

Increasing the cycle length at a single intersection tends to disrupt progression at adjacent intersections, so there is a limit to the amount of time that should be added to a cycle. One exception would be when the downstream intersection is capable of handling the increased traffic. In addition to typical cycle length factors (e.g., volume and capacity), oversaturated conditions require consideration of factors such as available storage space, arrival rates on red, and split ratios. A cycle length policy developed by Lieberman, Chang, and Prassas (*2*) puts an upper bound on cycle lengths to avoid spillback and resulting progression issues. These cycle lengths ensure that the queue formation shockwave dissipates before reaching the upstream intersection.

An example based on the Lieberman, Chang, and Prassas policy shows a family of curves (illustrated in Exhibit 12-10) that relate the highest feasible cycle length to specific link lengths and split ratios. The blue lines in the exhibit are the upper bound for cycle lengths, and the area beneath them represents the cycle lengths that will not generate spillback. For example, a split ratio of 0.5 for the downstream through phase and a 700-foot link length yields a maximum cycle length of 150 seconds before spillback occurs. This graph should be viewed as a rough guide when selecting cycle lengths; many other factors must be considered, including demand, offset, pedestrian requirements, and minimum green times.

Substantial increases or decreases in cycle length can cause excessive durations of transition (introduced in Chapter 7), and the effectiveness of interim signal timing parameters during transition periods can be poor. In particular, if the minimum green time for a phase with an overflow queue is reduced during transition, the transition period will only worsen conditions. For the mitigation of oversaturated conditions, it is important that any new cycle length is implemented before queues begin to form.

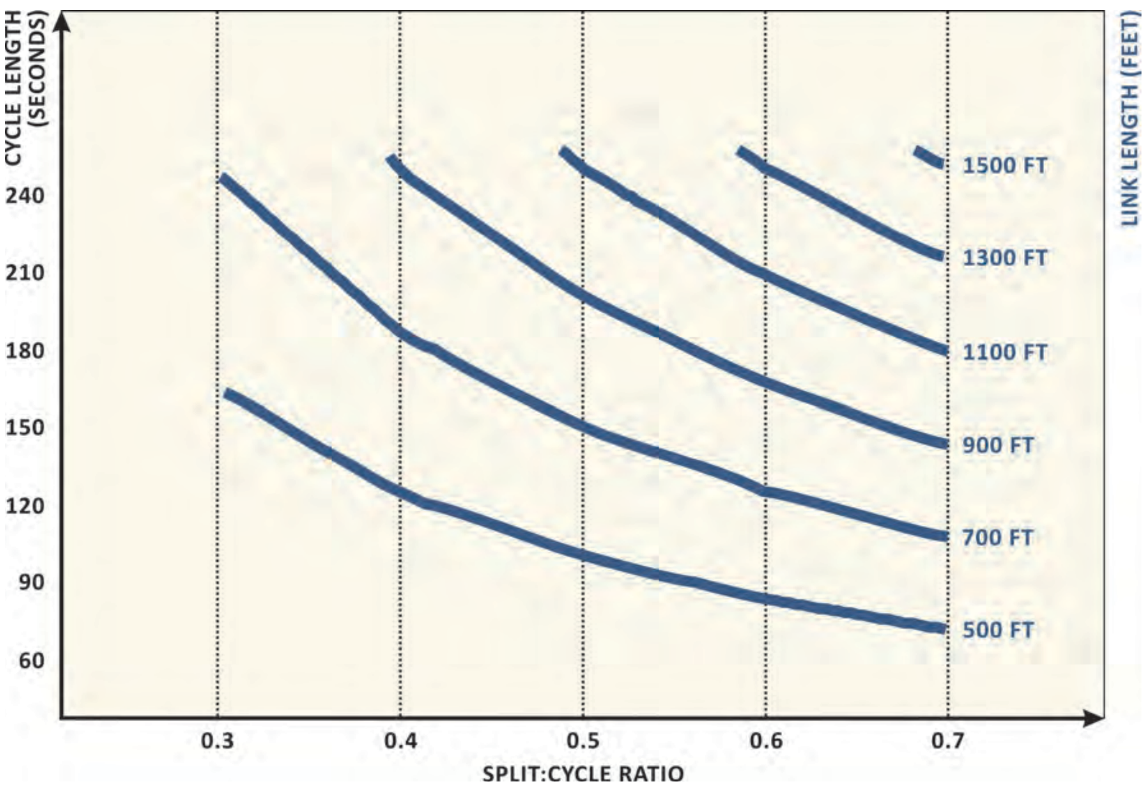


Exhibit 12-10 Example of Cycle Length as a Function of Split Ratio and Link Distance

12.3.1.3 Operation of Closely Spaced Intersections on One Controller

Diamond interchanges are the most common example of locations where more than one intersection may operate jointly on one controller. Such operation permits close coordination between intersections, which can help minimize spillback and starvation, increase system throughput, and manage queues on links with limited storage space. The operation of closely spaced intersections on one controller requires the strategic placement of field detection (if actuated), as well as the installation of cabinet wiring specific to this operational structure. Thus, this strategy is not intended to be applied as part of a time-of-day schedule or as an intermittent, reactive solution to traffic conditions.

If two intersections are fewer than 10 seconds of travel time apart, the use of one controller for intersection operations should be considered. (At a travel speed of 35 miles per hour, this distance is approximately 500 feet.) Most modern controllers can accommodate this type of operation through the use of overlaps and careful setup of phase sequence and splits. Operation of two intersections on one controller may necessitate the use of more than eight phases, four overlaps, and two rings. Several traffic control firmware programs and state-of-the-art cabinets are available that can accommodate this specialized operation.

12.3.1.4 Green Extension

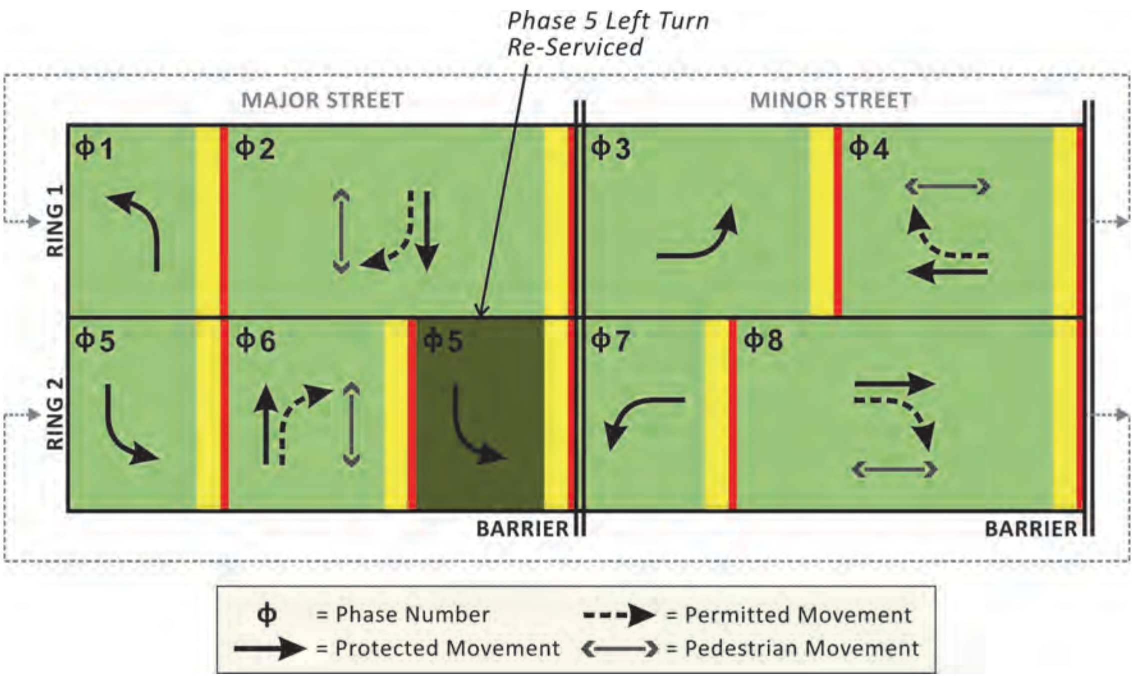
Green extension is a strategy that adds green time to a phase when a detector (or detectors) on an approach exceed some threshold of occupancy. This strategy is often applied at freeway-to-arterial interchanges where off-ramp queues may affect freeway flow and pose safety concerns. It is typically enabled through logic in the local controller that activates an alternative split or, in extreme cases, through a preemption input that immediately serves the oversaturated phase for a predetermined amount of time.

Despite its utility, preemption disrupts coordinated operations and is not a recommended strategy to mitigate oversaturation. If the strategy is applied, however, care should be taken to consider the storage space available downstream. If limited storage is available, oversaturation should be addressed across the span of the route as a systemic issue, not as an isolated issue at a single intersection.

12.3.1.5 Phase Re-Service

Phase re-service is a strategy that serves a phase twice during the same cycle. Intersections that benefit from phase re-service are often characterized by heavily imbalanced flows on major movements. A commonly implemented phase re-service scenario is one in which a left-turn phase both leads and lags the opposing through movement (illustrated in Exhibit 12-11). In a modern controller, this is configured through activation of a “phase re-service” or “conditional service” setting.

Exhibit 12-11
Ring-and-Barrier Diagram for Re-Service of Left-Turn Phase



Another scenario in which phase re-service may be effective is when an intersection approach is congested due to the merging of two heavy movements upstream of the intersection. For example, a shared receiving lane for an upstream southbound left-turn movement and eastbound through movement (as shown in Exhibit 12-12) may cause congestion downstream. In this situation, traffic volumes approach the downstream intersection at two distinct times during the cycle. Excessive delay for one of the two upstream movements may result if the phase at the downstream intersection is only served once per cycle.

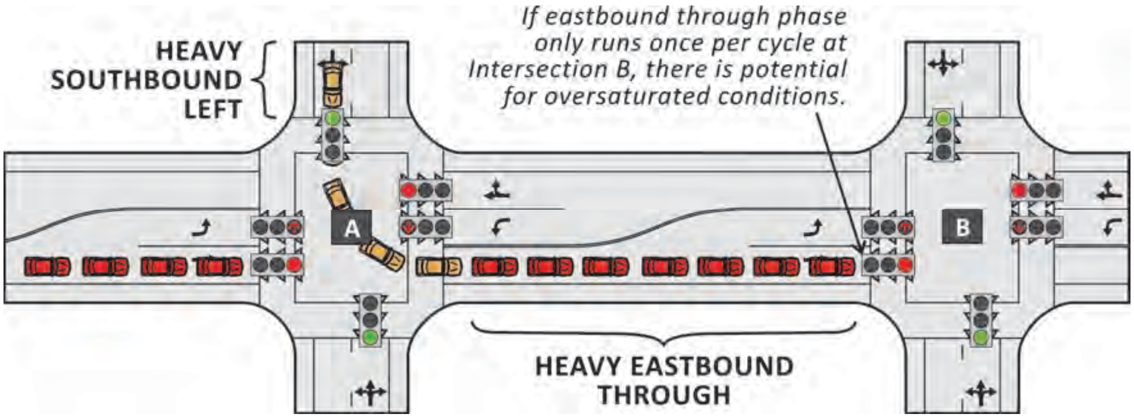


Exhibit 12-12 Merging of Two Heavy Upstream Movements

Another application of the phase re-service strategy is alternating service of the minor movement every other cycle. This strategy may be useful for intersections that have split phasing on the minor street, a constrained cycle length, and are unable to serve both minor street approaches while providing enough time to the major street phases. To alleviate this problem, one minor street approach may be served at the start of the cycle, while the other minor street approach is served toward the cycle's end. This strategy is also considered to be "double cycling," as all phases are served only once in twice the amount of time as those at adjacent intersections. While there are traditionally eight phases in a dual-ring controller configuration, double cycling requires the use of overlaps. Most modern controllers with sixteen phases and eight overlaps can be configured to provide this type of operation (as illustrated in Exhibit 12-13).

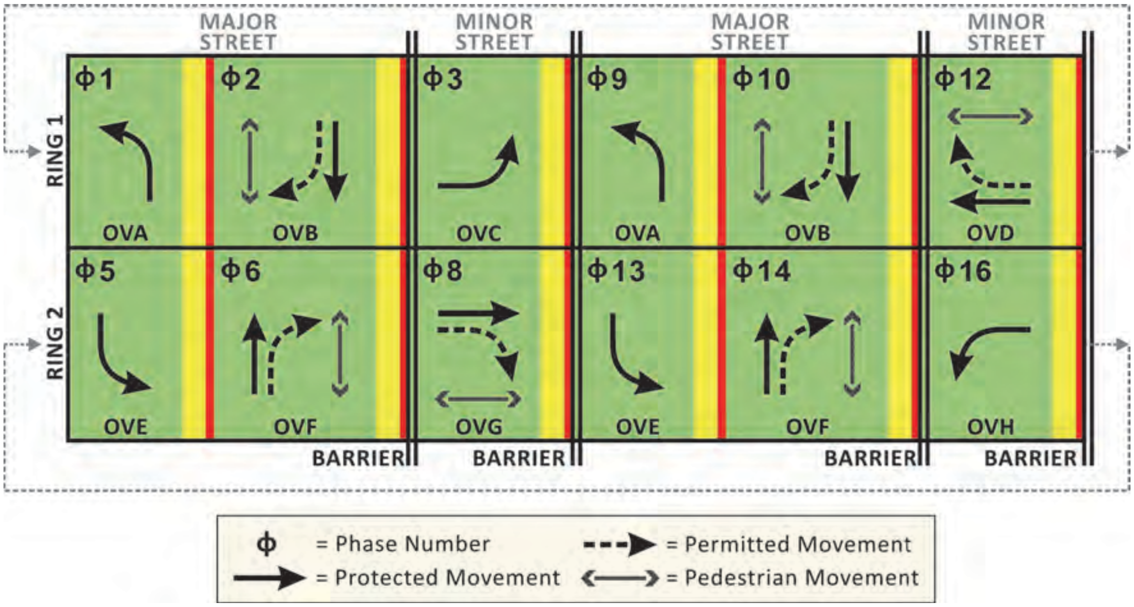
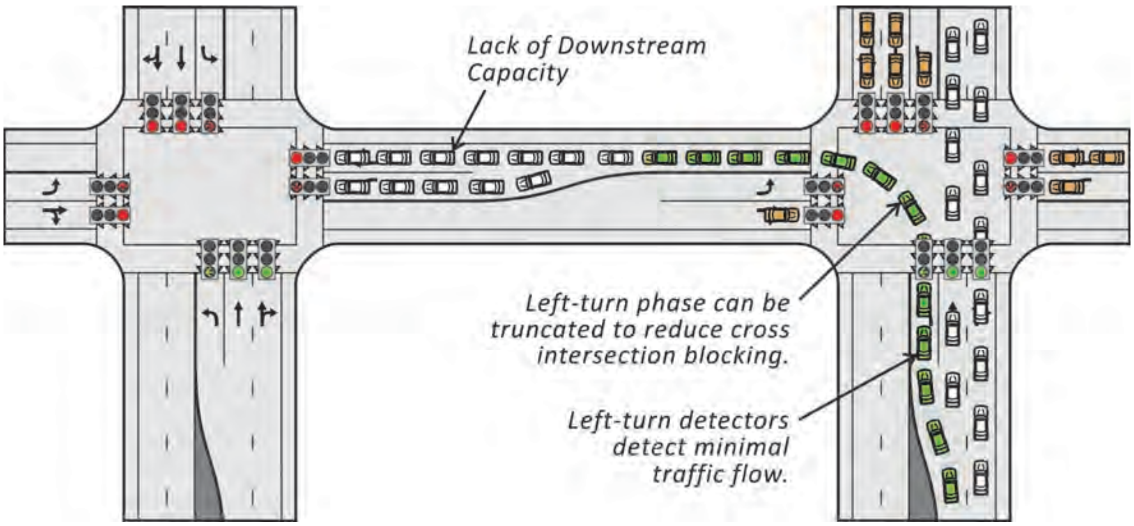


Exhibit 12-13 Ring-and-Barrier Diagram for Double Cycling

12.3.1.6 Phase Truncation

Phase truncation is a strategy that terminates the green interval (despite demand) when there is minimal to no flow over a detector, using logic programmed in the local controller. This strategy may be applied at intersections where there is a lack of downstream capacity or where overflow queues for other phases are common, as illustrated in Exhibit 12-14. In the exhibit, northbound left-turning vehicles block the intersection due to downstream congestion. Truncation of the northbound left-turn phase prior to the resultant cross blocking permits other movements to be served while the downstream queue disperses. When a phase is truncated, the spare green time is reallocated to provide extended service to other phases.

Exhibit 12-14 Phase Truncation Example



This type of mitigation is most applicable when queue management is the primary objective and when intermittent oversaturation occurs downstream of the affected phase. Research for NCHRP Project 03-66 (*3*) concluded that, in an isolated situation, phase truncation can effectively reduce lost green time. However, the effectiveness of truncation depends on the growth and dissipation patterns of the overflow queue. For example, if left-turn demand is relatively low and left-turning vehicles do not spillback into the through lane, the overall operational improvements are likely to be minor.

12.3.2 Mitigation Strategies for Arterials

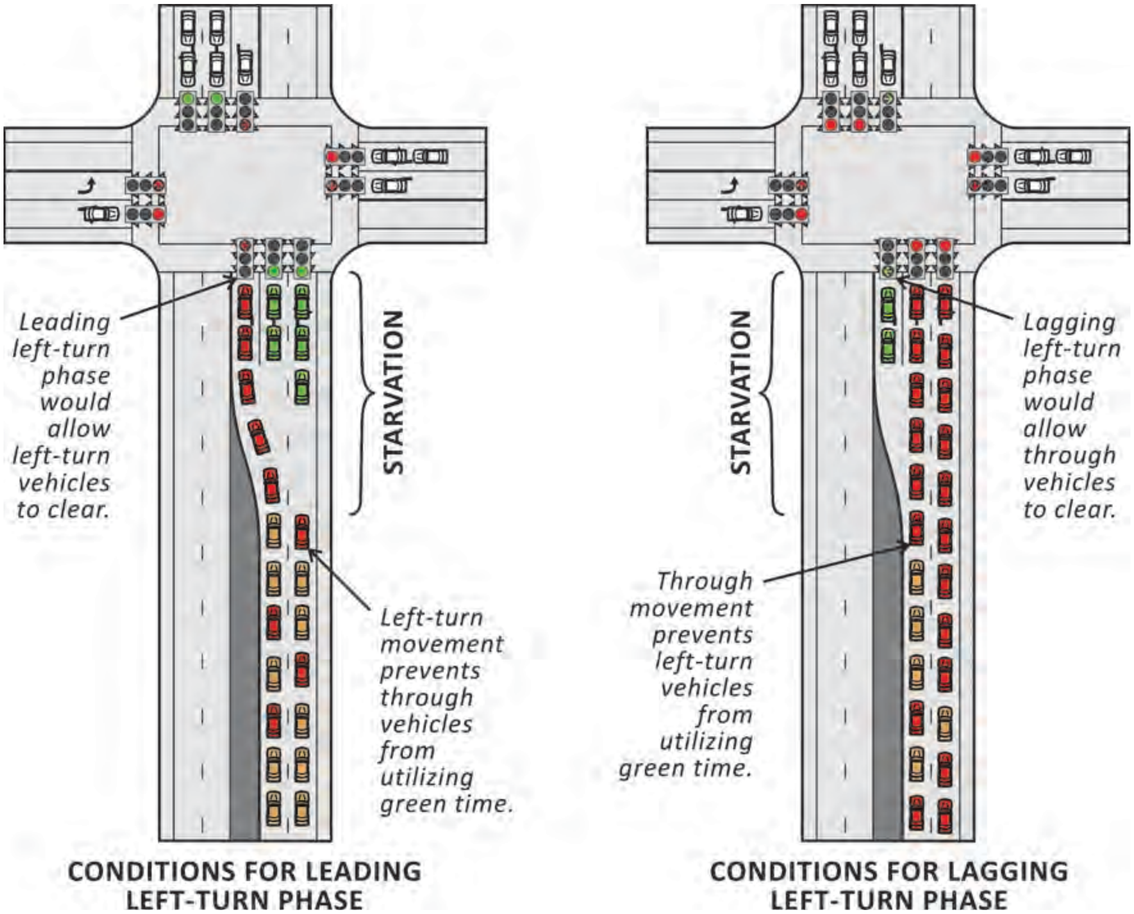
In most cases, using mitigation strategies for oversaturated arterials in combination with intersection-specific strategies is recommended.

12.3.2.1 Phase Sequence Modification for Left Turns

Oversaturated conditions can be addressed by modifying the left-turn phase order. In particular, using lead-lag left-turn phasing provides a wider bandwidth for the progression of platoons (introduced in Chapter 7). In oversaturated conditions, lead-lag phasing can help prevent storage bay spillback, storage bay blocking, and starvation. While these principles generally apply to arterial oversaturation, individual intersections and minor movements may benefit as well. The alteration of lead-lag settings at an intersection is not a strategy that can typically be applied for short-term, intermittent conditions. When the phase sequence is altered, modern controllers induce transition to reestablish a new offset reference point. If permitted left-turn phasing is used, flashing yellow arrow (FYA) operation is necessary to accommodate the phase sequence switch.

The overall corridor timing plan, as well as the individual intersection conditions, will dictate which phase to lead and which to lag in a lead-lag sequence (as shown in Exhibit 12-15). If a left-turn movement at a particular approach frequently overflows the available storage bay, the left-turn movement should lead the through movement in order to clear left-turning vehicles. Conversely, a lagging left turn is more appropriate if the through lane typically backs up past the end of the left-turn storage bay. This phase order permits the through vehicles to proceed through the intersection and allows the trapped left-turn vehicles to enter the left-turn lane prior to the beginning of the left-turn phase. For a lagging left turn, additional controller programming is recommended, so that the phase can gap out (and not extend unnecessarily to the end of the coordinated through phase). However, because lagging left turns can use the full split time, it may be beneficial to lag the left-turn movement with the higher demand.

Exhibit 12-15
Conditions for Leading and Lagging Left-Turn Phases



12.3.2.2 Flushing Strategies

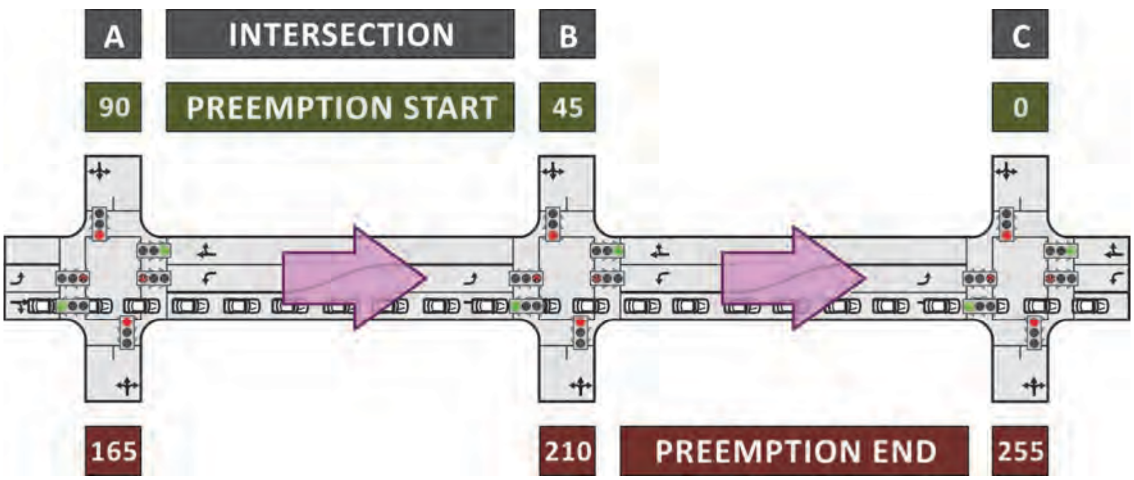
Overly congested, persistent queuing on an oversaturated route is difficult to mitigate with measures such as offset adjustment or phase sequence modification. In these extreme conditions, preemption (or “green flush”) strategies may be applied after other mitigation strategies have provided little incremental benefit. In conditions where minor street volumes compose a notably lower proportion of the total system traffic (i.e., 20 percent of total system volume), the effectiveness of flushing strategies is substantially greater than that of other mitigation strategy combinations. If minor street volumes compose a larger portion of system traffic, however, application of green flush strategies will considerably increase minor street delay.

There are two primary methods used to flush congested corridors. One method utilizes a series of preemption calls to controllers on the route to achieve vehicle progression. The second method initiates a timing plan with a long cycle length that allocates a majority of the split time to the oversaturated route. Either approach can be applied offline by time-of-day plans, operators at a traffic management center (TMC), or via logic. Operators in Sacramento County, California, for example, manually initiate a route preemption strategy from a TMC based on closed-circuit television (CCTV) surveillance.

12.3.2.2.1 Preemption or "Green Flush"

In a green flush operation, a series of preemption calls are initiated in sequence from the most downstream location to the most upstream location (as illustrated in Exhibit 12-16). This differs from emergency service preemption, which places calls in the direction of travel. This "backward" sequence aims to clear downstream queues before the release of traffic at upstream signals. For effective implementation of the preemption strategy, the downstream preemption must last as long as necessary for the traffic from upstream locations to progress through the downstream location (or as close as possible, with the 255-second preemption time limitation in most field controllers). Depending on the persistence of arrival volumes on the oversaturated route, it may be necessary to periodically apply route preemption as queues rebuild during "normal" intersection operations.

Exhibit 12-16 Example Green Flush Sequence



One challenge with using a preemption strategy is the adverse effect of post-preemption transition. Before returning to normal operations, most modern controllers can be configured to serve the minor street phases immediately after preemption ends. This exit behavior should be carefully considered when applying a green flush strategy.

12.3.2.2.2 Flushing through an Alternative Timing Plan

The less severe alternative to the green flush strategy is to implement a timing plan with a long cycle length that allocates a majority of the split time to the oversaturated route. This approach periodically allows a small window of travel opportunity for under-saturated movements, although a majority of the cycle is dedicated to progression on the oversaturated route. This method often results in significant queuing on minor streets and under-saturated movements, but can be scheduled as a time-of-day pattern if route oversaturation is recurrent and predictable. As shown in Exhibit 12-17, this flush strategy significantly improves the travel time of the oversaturated route when compared to less aggressive mitigation strategies.

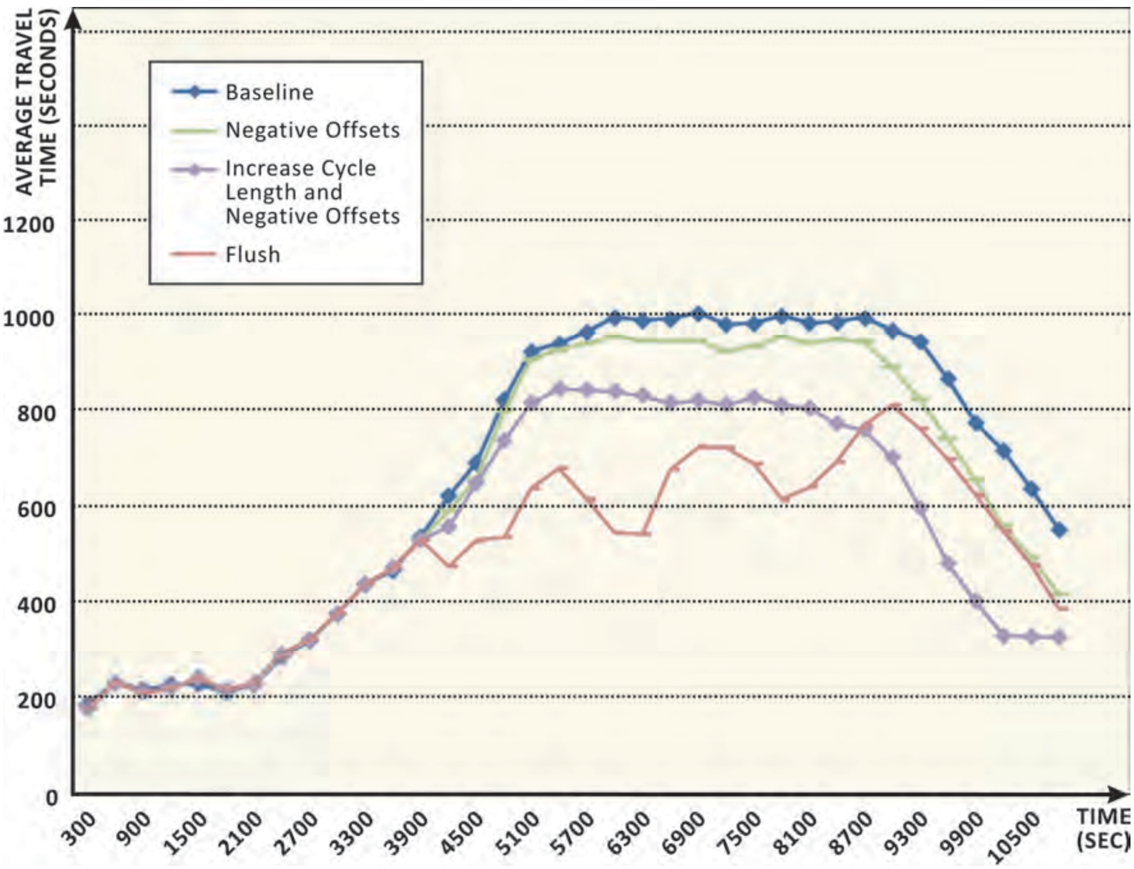


Exhibit 12-17 Average Travel Time Using Different Oversaturation Mitigation Strategies

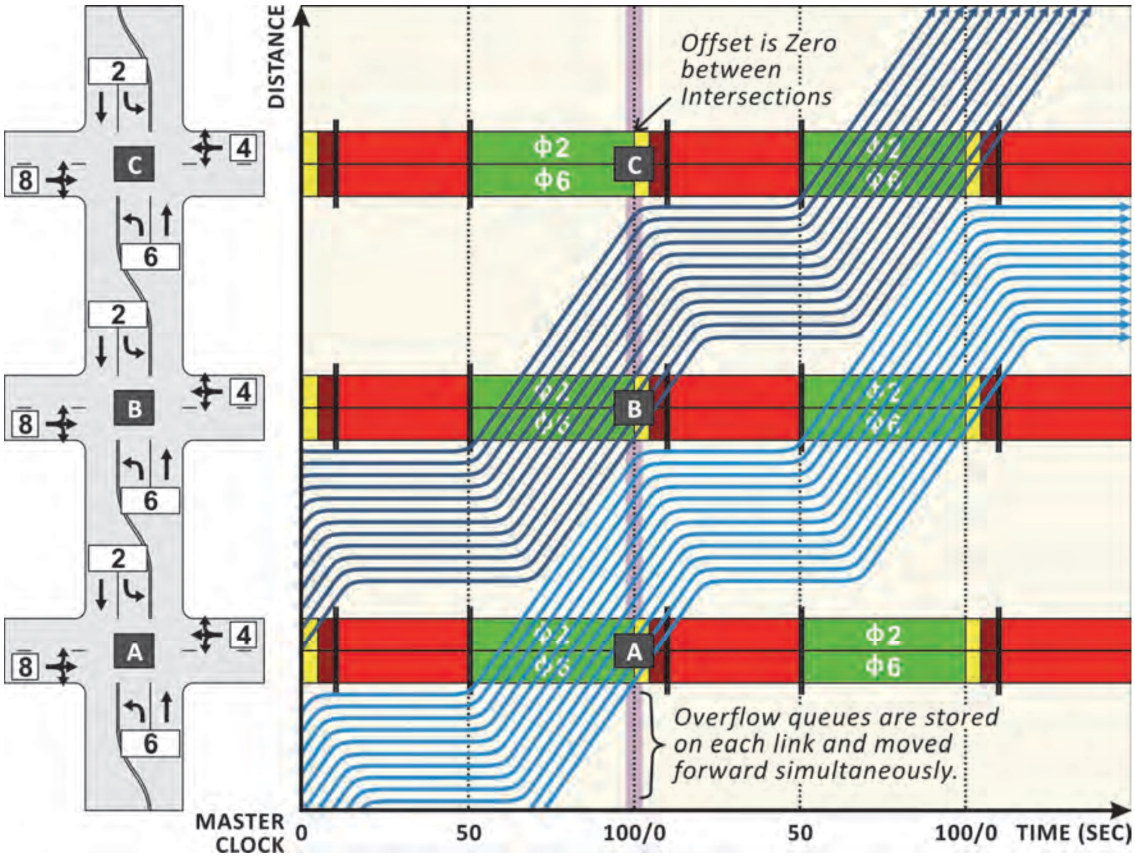
12.3.2.3 Offset Strategies

Offsets establish green bands for efficient progression along an arterial (see Chapter 7 for more information on bandwidth). Most algorithms used to design green bands assume that vehicles can progress along a route unimpeded by overflow queues, and the majority of offset design algorithms assume that queues clear before the first platoon arrives at the downstream stop bar. When overflow queues begin to form, however, these assumptions are violated, and the practitioner must take a closer look at appropriate offsets.

12.3.2.3.1 Simultaneous Offsets

A simultaneous offset is a special case in which the relative offset between two intersections is set to zero (illustrated in Exhibit 12-18). Applications of simultaneous offsets are most effective at locations where the overflow queue on each approach is approximately the same, the green time of each coordinated phase is approximately equal, and, therefore, the queue growth rate is approximately the same. This approach is sometimes referred to as “store and forward.”

Exhibit 12-18 Example of Simultaneous Offsets



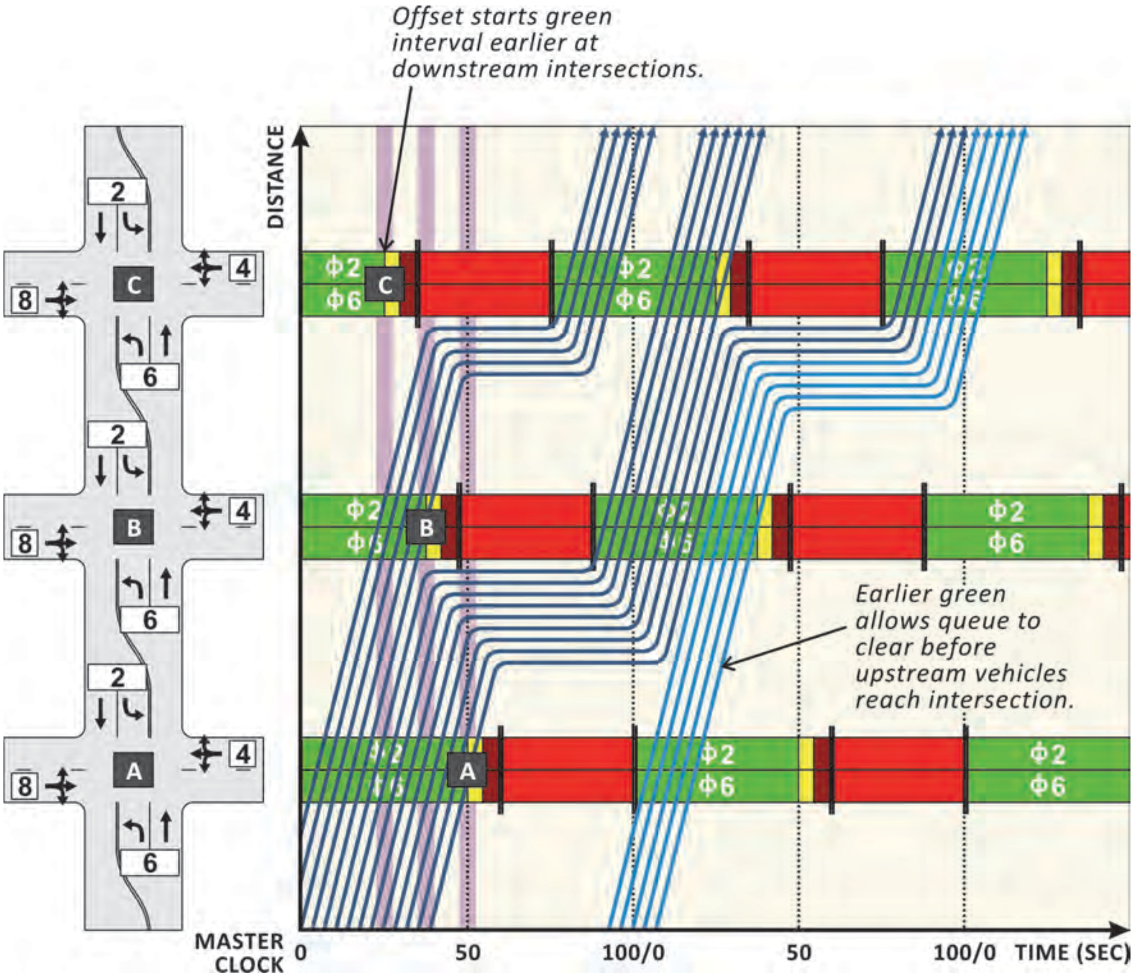
12.3.2.3.2 Negative Offsets

Negative offsets are characterized by starting the green interval earlier at the downstream intersection than at the upstream intersection (as illustrated in Exhibit 12-19). This allows the downstream overflow queue to disperse before the arrival of the upstream platoon and mitigates wasted green time that results when an upstream platoon arrives at the back of a downstream queue. Generally, the shorter the distance between two intersections, the more critical the design of offsets becomes.

Negative offsets are particularly effective with directional traffic, while simultaneous offsets may be a more practical solution when congestion exists in both directions. As previously mentioned, more efficient operations may be realized with a combination of strategies. Green reallocation and cycle length adjustments coupled with negative (or simultaneous) offsets can improve progression and minimize lost time. Traffic demand at downstream intersections can also be reduced by storing traffic upstream; metering is discussed in detail in Section 12.3.3.1.

Negative offsets alone will not clear an overflow queue if the green time is not adequate for the traffic demand, but they may reduce the adverse effects of positive progression offsets designed without consideration for downstream queues. Once a queue grows to the extent that all green time for the phase is allocated solely to overflow queue clearance (i.e., vehicles in the queue take more than two cycles to clear the intersection), the relationship between offsets is no longer meaningful. Without an increase in green time for the oversaturated phase, adjustments to negative offsets can only marginally reduce the growth rate of overflow queues. Once heavy demand rates subside, however, negative offsets are more efficient than positive offsets at dispersing overflow queues formed during oversaturation.

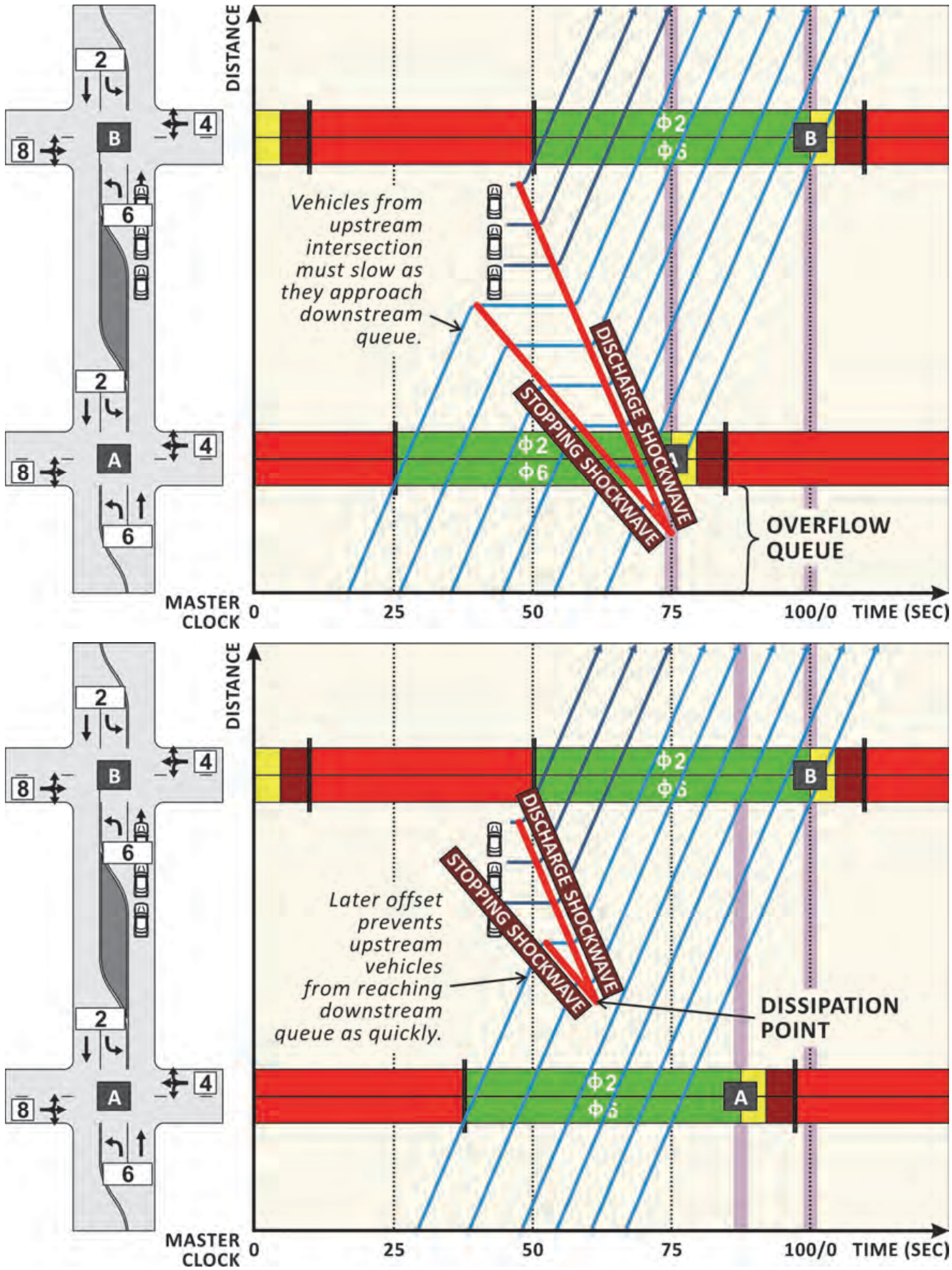
Exhibit 12-19 Example of Negative Offsets



12.3.2.3.3 Offsets to Prevent Queue Spillback

Queue spillback forms when a downstream queue fills the downstream link, preventing upstream vehicles from proceeding through the intersection. When a queue exists at a downstream intersection, vehicles must slow in order to join the back of the queue (as shown in Exhibit 12-20). Particularly for systems with very short link lengths (i.e., in downtown areas), it is important to design the progression strategy appropriately to prevent a ripple effect. Offset settings can help prevent queue growth by providing additional time for a downstream queue to clear before the upstream vehicles reach the end of the downstream queue.

Exhibit 12-20 Example of Offset to Prevent Queue Spillback



12.3.2.3.4 Offsets to Prevent Starvation

Starvation occurs when vehicles discharging at an upstream intersection arrive later than the time when the standing queue at a downstream intersection has been discharged. Starvation results in loss of capacity by wasting valuable green time. An ideal offset ensures that the first released vehicle at the upstream intersection joins the discharging queue at the downstream intersection just as the back of the queue begins to move. Any further delay in the release of vehicles from the upstream location may cause downstream starvation (illustrated in Exhibit 12-21).

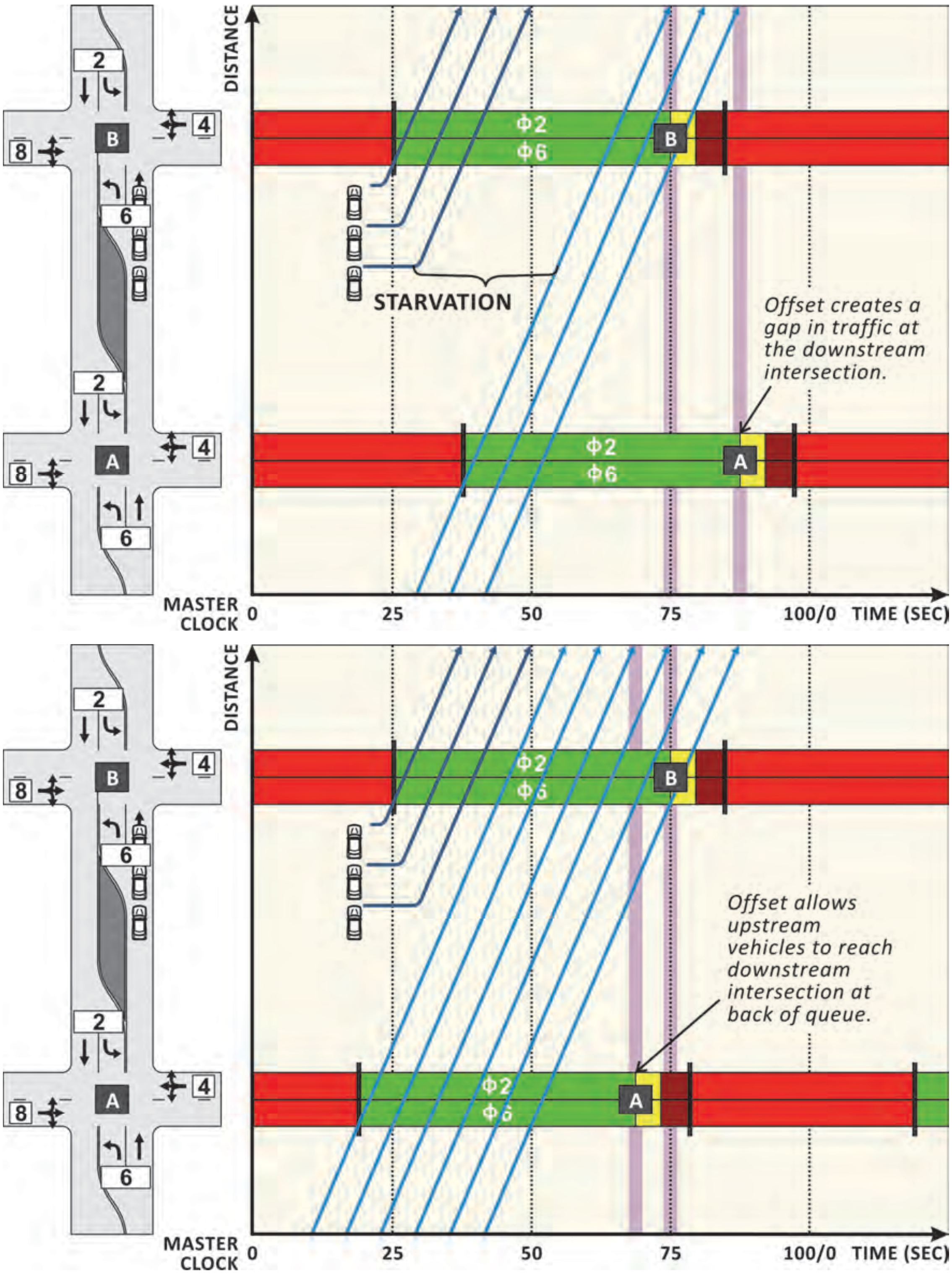
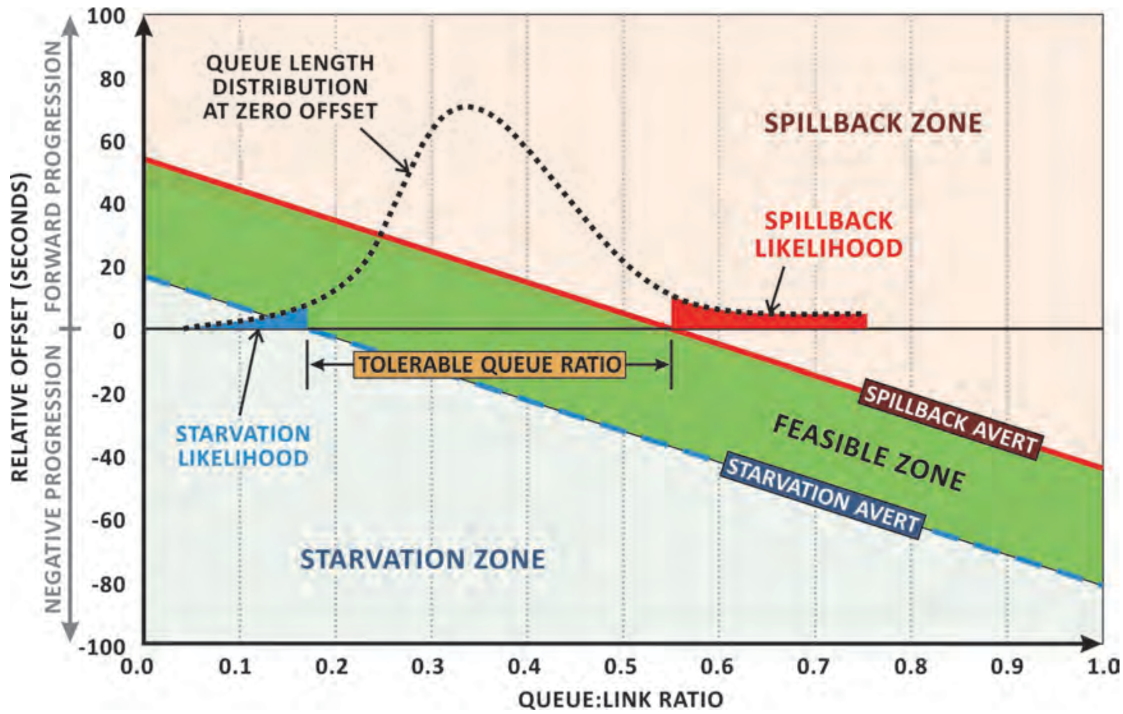


Exhibit 12-21 Example of Offset to Prevent Starvation

12.3.2.3.5 Combining Offset Strategies

Exhibit 12-22 illustrates how various strategies can be used to determine a desirable range of offsets when queues are present. The range of desirable offsets depends on link length, overflow queue length, and queue discharge rates; the resulting feasible zone is defined using relative offsets. For example, a queue ratio of 0.5 (i.e., half of the 800-foot link length is filled with a queue) constrains the feasible offsets between relative negative 30 seconds and 10 seconds. The larger the queue, the larger the need for negative relative offsets to avoid downstream congestion. As discussed previously, a negative relative offset may be infeasible with short block spacing and the need to move congested traffic in both directions.

Exhibit 12-22 Offset Values to Avoid Spillback and Starvation for Intersections Spaced at 800 Feet



12.3.3 Mitigation Strategies for Networks

As oversaturated conditions worsen and affect more intersections and arterials, a more extensive, network-level approach may be necessary.

12.3.3.1 Metering (Gating)

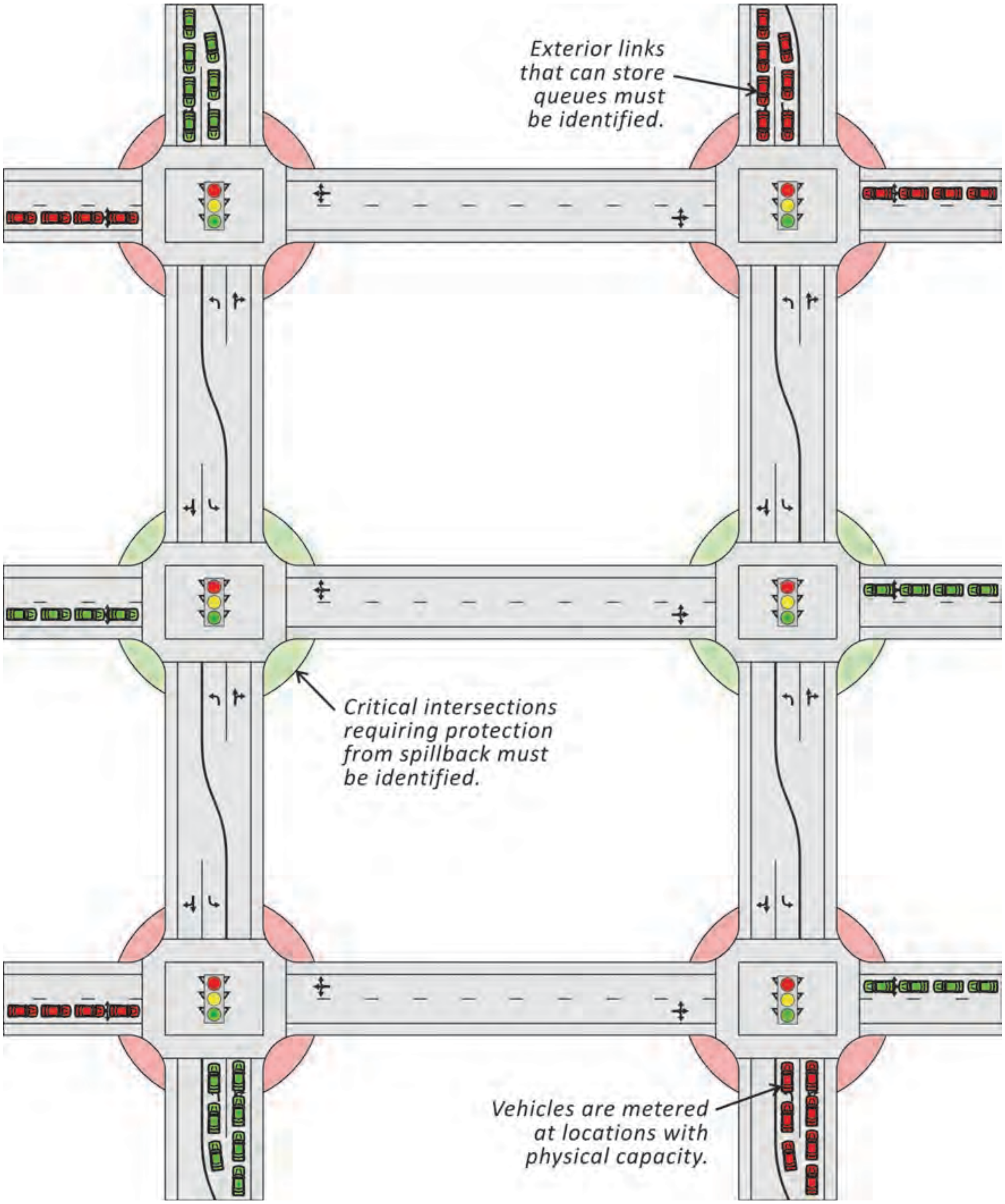
Metering (also known as gating) can be applied to impede traffic at appropriate upstream points and to prevent traffic flow from reaching critical levels at downstream intersections. Application of the gating strategy is appropriate when downstream spillback is the main cause of congestion. Implementation of this strategy involves identification of critical intersections in the network that require protection from spillback, as well as identification of exterior links that can be used to store queues. Metering locations should be determined on a case-by-case basis; general considerations for this selection include

- Physical link length and storage capacity,
- Locations of traffic generators and traffic attractors, and

- Access to generators and attractors.

Exhibit 12-23 illustrates an application of gating. In the exhibit, the depicted oversaturated network can be alleviated with the application of a gating strategy at several points (vehicles shown in red). By reducing green time on those approaches, overflow queues can be contained at the network entrance points. However, care must be taken to prevent the oversaturation problem from shifting to another network outside of the consideration area. If appropriately applied, gating can also be used in conjunction with other mitigation strategies to improve performance.

Exhibit 12-23 Example of Gating Operation



During service recovery, gating is typically lifted so that spare network capacity can be used and queues formed outside of the network can be alleviated. If oversaturation is predictable and recurrent, a time-of-day schedule may be developed that applies these strategies during each oversaturation performance period. If the start and end times of network saturation are variable, the decision to use gating may be executed through logic statements or manual control (e.g., TMC operator control).

12.3.4 Adaptive Control

Adaptive control systems should not be expected to eliminate oversaturation. Rather, these systems are one tool in a box of many other conventional strategies that address the issue of oversaturated conditions.

Adaptive control methods (described in Chapter 9) can improve performance over time-of-day and actuated control operations in under-saturated conditions. While adaptive methods can be effective in delaying the onset of congestion, the ability of adaptive systems to mitigate oversaturated conditions has not been demonstrated. The core features of adaptive control that help to delay the onset of oversaturation include cycle length adjustment, split reallocation, offset adjustment, and phase sequence modification. If congestion is extensive in time and space, the previously described techniques are better able to manage congestion and increase throughput. If congestion is minor, adaptive systems can recover from oversaturation faster than conventional controls.

12.4 REFERENCES

1. Denney, Jr., R. W., L. Head, and K. Spencer. *Signal Timing Under Saturated Conditions.* Report FHWA-HOP-09-008, Federal Highway Administration, United States Department of Transportation, 2008.
2. Lieberman, E. B., J. Chang, and E. S. Prassas. Formulation of Real-Time Control Policy for Oversaturated Arterials. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1727*, Transportation Research Board, National Research Council, Washington, D.C., 2000, pp. 77–88.
3. Beaird, S., T. Urbanik, and D. M. Bullock. Traffic Signal Phase Truncation in Event of Traffic Flow Restriction. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1978*, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 87–94.

APPENDIX A. GLOSSARY

The following is a collection of terms used in this edition of the *Signal Timing Manual*. Some terms may have different definitions in other references, but the definitions below are consistent with the usage of terms in this manual.

Actuated Signal Control—Phase time based on detection. *See also Fully-Actuated Control and Semi-Actuated Control.*

Adaptive Signal Control—An advanced signal system that does not operate with time-of-day plans.

Analysis Period (or Time Interval)—A single time period during which capacity analysis is performed on a transportation facility. If the demand exceeds capacity during an analysis period, consecutive analysis periods can be selected to account for the initial queue from the previous analysis period.

Analytical Model—A model that relates system components using theoretical considerations (tempered, validated, and calibrated by field data).

Approach—A set of lanes at an intersection that accommodates all left-turn, through, and right-turn movements from a given direction.

Approach Grade—The grade of an intersection approach, expressed as a percentage (with positive values for upgrade and negative for downgrade).

Arterial—A signalized street that primarily serves through traffic and that secondarily provides access to abutting properties, with signals spaced 2 miles (or less) apart.

Average Speed—The average distance a vehicle travels within a measured amount of time. Typically, average speed is measured over a short distance. *See also Average Travel Speed.*

Average Travel Speed—The length of the highway segment divided by the average travel time of all vehicles traversing the segment, including all stopped delay times. *See also Average Speed.*

Back of Queue—The distance between the stop line of a signalized intersection and the farthest reach of an upstream queue, expressed as a number of vehicles. The vehicles previously stopped at the front of the queue are counted even if they begin moving. *See also Queue.*

Bandwidth—The maximum amount of green time for a designated coordinated movement as it passes through a corridor at an assumed constant speed, typically measured in seconds.

Barnes' Dance—A common term for an exclusive pedestrian phase where pedestrians may cross all intersection legs (and sometimes diagonally). *See also Pedestrian Phase.*

Barrier—A reference point in the sequence where two or more rings are interlocked. Barriers ensure that there will be no concurrent selection and timing of conflicting movements in different rings.

Bike—A vehicle type which may require special timing consideration.

Cabinet—An all-weather enclosure that houses the intersection field equipment.

Call—A request for service. Some movements may have separate calls for different user types. For example, a transit priority call is one where the normal phase timing may be altered due to a call from a transit vehicle. More typical calls are vehicle calls and pedestrian calls.

Capacity—The maximum rate at which vehicles can pass through an intersection under prevailing conditions.

Change Interval—The yellow change interval plus red clearance interval that occurs between phases of a traffic signal in order to clear the intersection before conflicting movements are released.

Concurrent Phases—Two or more phases in separate rings that are able to operate together without conflicting movements.

Congested Flow—A traffic-flow condition caused by a downstream bottleneck.

Controller—The piece of hardware that determines how a traffic signal responds to calls based on signal timing parameters.

Controller Memory—A term that refers to a controller's ability to "remember" (i.e., retain) a detector actuation or not remember a detector actuation. There are two modes (non-locking and locking). The locking mode remembers the actuation after it is dropped by the detection unit; the non-locking mode does not remember the actuation.

Coordinated Phase(s)—The phase (or phases) that are given a fixed minimum amount of time each cycle under a coordinated timing plan. This phase is typically the major through phase on an arterial. Coordinated phase(s) may also have an optional actuated interval following the fixed interval.

Coordination (or Coordinated)—The ability to synchronize multiple intersections to enhance the operation of one or more directional movements in a system.

Corridor—A set of essentially parallel transportation facilities designed for travel between two points. A corridor contains several subsystems, such as freeways, arterials, transit, and pedestrian and bicycle facilities.

Critical Movement Analysis—A simplified technique used to identify the critical movements at an intersection, estimate whether the intersection is operating adequately, and approximate the amount of green time needed for each critical movement.

Crosswalk—A marked area for pedestrians crossing the street at an intersection or designated midblock location.

Cycle Length—The duration of a complete sequence of phases in the absence of priority calls. In an actuated controller unit, a complete cycle is dependent on the presence of calls for all non-priority phases. Some indications may be served more than once in a cycle. Occasionally, an indication may not be part of a normal cycle (e.g., a left-turn arrow may only be displayed during railroad preemption).

Decision Zone—An area in front of a stop bar where some drivers would choose to stop and others would choose to proceed through the intersection, upon the change from a green to yellow indication. Detection designs can be used to reduce the probability of drivers having to make this type of decision on high-speed approaches. This area is also known as option zone, indecision zone, and Type II dilemma zone. The term decision zone is used in this manual to avoid confusion with dilemma zones that occur due to improper clearance. *See also Dilemma Zone.*

Delay—(1) The additional travel time experienced by a driver, passenger, or pedestrian. (2) A detector parameter typically used with stop-bar presence-mode detection for turn movements from exclusive lanes.

Demand—The number of users desiring to use an intersection, approach, or movement during a particular time period. Not to be confused with volume, which is a measure of the number of users accommodated at the intersection (which is limited to the available capacity).

Density—The number of vehicles on a roadway segment averaged over space, usually expressed as vehicles per mile or vehicles per mile per lane. *See also Volume-Density*.

Detector (or Detection)—A device used to count and/or determine the presence of a motorized vehicle, bicycle, or pedestrian. *See also Setback Detection and Stop-Bar Detection.*

Detector Card(s)—The detector processor module that is installed in a detector rack within a cabinet.

Detector Delay—*See Delay.*

Detector Extend—*See Extend.*

Detector Rack—The hardware module that holds detector cards within a cabinet.

Detector Switching—A common detector function in traffic signal controllers that allows detectors to extend calls for one phase (extend phase) and then send calls to another phase (switch phase) once the extend phase ends. Detector switching allows the programmed switch phase to be extended after the extend phase terminates. It is typically only effective when the switch phase is green. Also, detector switching does not typically switch phase calls.

Dilemma Zone—A condition that occurs when yellow change and red clearance times are too short for a driver to either stop or clear the intersection before the beginning of a conflicting phase. Also known as Type I dilemma zone. *See also Decision Zone.*

Display (or Head, Signal Group)—A combination of indications (e.g., red, yellow, green, green arrow, audible) grouped together for controlling one or more movements.

Double Cycle—A cycle length that allows phases at an intersection to be served twice as often as the phases at other intersections in the coordinated system.

Downstream—The direction of traffic flow. Generally the discharge side of an intersection.

Dual Entry—The parameter used to call vehicle phases that can time concurrently, even if only one of the phases is receiving an active call. For example, if dual entry is active for Phases 4 and 8, and Phase 4 receives a call but no call is placed on Phase 8, Phase 8 will still be displayed along with Phase 4. The most common use of dual entry is to activate the parameter for compatible through movements.

Early Return to Green—A term used to describe the servicing of a coordinated phase in advance of its programmed begin time as a result of unused time from uncoordinated phases.

Effective Green Time—The time during which a given traffic movement (or set of movements) may proceed; it is equal to the cycle length minus the effective red time. In a practical sense, effective green time is equal to actual green time, as the start-up lost time is approximately equal to the amount of time during the yellow change interval when vehicles are still entering the intersection.

Effective Red Time—The time during which a given traffic movement (or set of movements) is not moving into the intersection; it is equal to the cycle length minus the effective green time.

Exclusive Pedestrian Phase—A separate phase that is configured such that no vehicular movements are served concurrently with pedestrian traffic. *See also Pedestrian Phase.*

Exclusive Turn Lane—A designated left- or right-turn lane (or lanes) used only by vehicles making those turns.

Extend—A detector parameter that extends a detector actuation by a programmable fixed amount of time. It is typically used with detection designs that combine multiple setback detectors for safe phase termination of high-speed intersection approaches or to provide lane-by-lane detection.

Firmware—The software embedded in the traffic signal controller that operates the traffic signal system. Features may vary by firmware version.

Fixed Force-Off—A mode of split management used with coordinated operations, where force-off points cannot move. Under this mode, uncoordinated phases can utilize unused time from previous phases. *See also Force-Off.*

Fixed Time Signal Control—*See Pretimed Signal Control.*

Flashing Don't Walk—An indication warning pedestrians that the walk indication has ended and the don't walk indication is underway. The pedestrian clearance interval may be longer than the flashing don't walk interval, as it can include the yellow change and red clearance times.

Flashing Yellow Arrow—A type of signal head display that reduces "yellow trap" problems by providing a permitted indication that operates concurrently with the opposing through movement. It is physically separated from the adjacent through movement, minimizing the association of the through yellow with the left-turn yellow.

Floating Car Method—A commonly employed technique for travel time runs, which requires the vehicle driver to "float" with the traffic stream while traveling at a speed that is representative of the other vehicles on the roadway (i.e., pass as many vehicles as pass the floating car).

Floating Force-Off—A force-off mode where force-off points can move depending on the demand of previous phases. Under this mode, uncoordinated phases are limited to their defined split times, and all unused time is dedicated to the coordinated phases. Essentially, the split time is treated as a maximum amount for the uncoordinated phases. *See also Force-Off.*

Flow Rate—The equivalent hourly rate at which vehicles, bicycles, or pedestrians pass a point on a lane, roadway, or other trafficway. Computed as the number of vehicles, bicycles, or pedestrians passing the point, divided by the time interval (usually less than 1 hour) in which they pass, and expressed as vehicles, bicycles, or pedestrians per hour.

Force-Off—A point within a cycle where a phase must end regardless of continued demand. These points in a coordinated cycle ensure that the coordinated phase begins with enough time to maintain the designated offset. However, force-offs cannot override clearance times. *See also Fixed Force-Off and Floating Force-Off.*

Free Flow—A flow of traffic unaffected by upstream or downstream conditions.

Free Operation—*See Uncoordinated.*

Fully-Actuated Control—A signal operation in which vehicle detectors on each approach to the intersection control the occurrence and length of every phase. *See also Actuated Signal Control.*

Gap—The time separation between vehicles (in seconds). In other words, the time required for the front bumper of the second of two successive vehicles to reach the ending point of the rear bumper of the first. It is less than the headway between vehicles.

Gap Out—The passage timer has expired (timed out).

Gap Reduction—This is a feature that reduces the passage time to a smaller value while the phase is active.

Green Time—The duration (in seconds) of the green indication for a given phase at a signalized intersection.

Hardware—The devices that physically operate the signal timing controls, including the controller, detectors, signal heads, and signal monitor.

Hardware-in-the-Loop (HITL)—A means of providing a direct linkage between simulation models and actual signal controllers.

Headway—(1) The time (in seconds) between two successive vehicles as they pass a point on the roadway, measured from the same common feature of both vehicles (e.g., the front axle or the front bumper). (2) The time (usually expressed in minutes) between the passing of the front ends of successive transit units (vehicles or trains) moving along the same lane or track (or other guideway) in the same direction.

Indication—*See Display.*

Individual Signalized Intersection—An intersection controlled by a traffic signal. They are sometimes operated as an “isolated intersection,” which refers to a mode of operation, rather than a spatial relationship. Individual intersections may also be described as “free” operation, which indicates they are not currently being coordinated. An individual intersection may also be operated as part of a coordinated system for all or part of a day.

Inhibit Max—A basic timing parameter that removes maximum green as a phase parameter during coordination and allows a phase to extend beyond its normal maximum green time.

Input(s)—The contact closure from detectors that tells the traffic signal controller of the presence of vehicles (in general or by type) and pedestrians.

Intersection Delay (Average)—The total additional travel time experienced by users (as a result of control measures and interactions with other users) divided by the volume departing from an intersection. *See also Delay.*

Interval—The duration of time during which traffic signal indications (e.g., red, yellow, green, and flashing don't walk) do not change state (i.e., red interval, yellow interval, green interval, and flashing don't walk interval).

Isolated Operation—An intersection that is not currently being operated as part of a coordinated system. Also known as free operation. *See also Uncoordinated.*

Lane Assignment—Movements that are permitted from a specific lane.

Lane Utilization—The distribution of vehicles among lanes when two or more lanes are available for a movement. When lane utilization is unequal due to traffic patterns upstream or downstream, additional green time may be necessary beyond what would be needed for uniform lane distribution.

Leading Pedestrian Interval—A pedestrian interval option that starts a few seconds before the adjacent through vehicular phase, allowing pedestrians to establish a presence in the crosswalk, and thereby reducing conflicts with turning vehicles.

Lead-Lag Left-Turn Phasing—A left-turn phase sequence where one left-turn movement begins with the adjacent through movement and the opposing left-turn movement begins at the end of the conflicting through movement. This option may create a "yellow trap" with permitted signal displays that do not use the flashing yellow arrow.

Load Switch (or Switch Pack)—A device that allows the controller, which operates in a 12/24-volt DC environment, to direct a 120-volt AC current to various signal displays.

Local Controller—*See Controller.*

Locking Mode—A controller memory mode used to trigger a call for service with the first actuation received by the controller during the red interval. Typically only used when there is no stop-bar detection. *See also Controller Memory.*

Lost Time—The time per signal cycle during which the intersection is effectively not used by any movement; this occurs during the yellow change and red clearance intervals (clearance lost time) and at the beginning of most phases (start-up lost time).

***Manual on Uniform Traffic Control Devices for Streets and Highways* (MUTCD)—**The MUTCD, published by the Federal Highway Administration (FHWA), provides the standards and guidance for installation and maintenance for traffic control devices on roadways.

Master Clock—The background timing mechanism within the controller logic to which each controller is referenced during coordinated operations.

Master Controller—An optional component of a signal system that facilitates coordination of a signal system with local controllers.

Max Out—The maximum green has been reached.

Maximum Green—The maximum length of time that a phase can be green in the presence of a conflicting call.

Maximum Initial—The maximum period of time for which the added initial can extend the initial green period. This cannot be less than the minimum green time.

Maximum Recall— A recall mode that places a continuous call on a phase. *See also Recall.*

Measure of Effectiveness (MOE)—A quantitative parameter indicating the performance of a transportation facility or service.

Memory Mode(s)—*See Controller Memory.*

Minimum Gap—The volume-density parameter that specifies a minimum time between detector actuations (that is applied with the gap reduction feature). *See also Gap Reduction.*

Minimum Green—The minimum length of time that a phase must be green. It should be set based on driver expectancy and the storage of vehicles between the setback detectors and the stop bar (if stop bar presence detection is not used).

Minimum Recall—A recall parameter that times the minimum green for a phase, regardless of the demand on that movement. *See also Recall.*

Modifier Phase—A phase associated with an overlap that causes the overlap to be red when the modifier phase is green. *See also Overlap.*

Movement—A term used to describe the user (e.g., vehicle or pedestrian) action taken at an intersection (e.g., vehicle turning movement or pedestrian crossing). Two different types of movements include those that have the right-of-way (protected/exclusive) and those that must yield (permitted/permissive), consistent with the rules of the road or the *Uniform Vehicle Code*. *See also Permitted Movement and Protected Movement.*

Movement Priority—Each movement may be assigned a relative or absolute priority depending on the operating environment and locally desired outcomes.

Non-Locking Mode—A controller memory mode that does not retain an actuation in the controller after the actuation is dropped by the detection unit. *See also Controller Memory.*

Occupancy—The time that a detector indicates a vehicle is present. May also be expressed as a percent.

Offset—The time relationship between the coordinated phase(s) based on the offset reference point and a defined master reference (i.e., master clock or sync pulse).

Offset Reference Point (or Coordination Point)—The defined point that creates an association between a signalized intersection and the master clock.

Operating Environment—An area with similar characteristics that would have similar signal timing objectives.

Operating System—The controller processor system on which the firmware runs.

Operational Objectives—The desired signal timing outcomes for each operating environment and user group.

Output—The voltage from a load switch that powers a signal indication.

Overflow Queue—Queued vehicles that are present at the beginning of green and that are not served during the green interval at a signalized intersection, so must be served by a subsequent cycle (or cycles).

Overlap—A timing process that provides a way to operate a particular movement with one or more phases. It is a separate output that can use special logic to improve operations.

Oversaturation—A traffic condition in which the arrival flow rate exceeds capacity.

Parent Phase—A phase (or phases) used to determine whether an overlap should be active. Parent phases may be in different rings and/or may be on different sides of a barrier. *See also Overlap.*

Passage Time (or Vehicle Interval, Vehicle Extension, Gap Time)—A phase timer that ends a phase when the time from the last detector output to the controller exceeds the timer setting.

Pedestrian—An individual traveling on foot.

Pedestrian Clearance Interval—The time provided for a pedestrian to cross the entire width of an intersection. This interval is longer than the flashing don't walk interval, as it can include the yellow change and red clearance intervals.

Pedestrian Phase—Time allocated to pedestrian traffic that is typically concurrent with compatible vehicular phase(s). *See also Barnes' Dance and Exclusive Pedestrian Phase.*

Pedestrian Recall—A recall mode where there is a continuous call for pedestrian service, resulting in the pedestrian walk and clearance intervals timing every cycle. *See also Recall.*

Pedestrian Scramble—*See Barnes' Dance.*

Pedestrian Walk Interval—An indication that allows pedestrians to begin crossing the intersection.

Pedestrian Walking Speed—The average walking speed of pedestrians (in feet per second).

Performance Measures—Quantifiable means that are used to assess whether a signal system has achieved its operational objectives. Examples include stops, arrival on green, phase failure, arterial travel time, vehicle delay, and existence of spillback queuing from turn bays or between closely spaced intersections.

Permissive Movement—*See Permitted Movement.*

Permissive Period—A period of time during the coordinated cycle in which calls on conflicting phases will result in the coordinated phase transitioning to an uncoordinated phase.

Permitted Movement—A movement that is allowed to proceed if there are available gaps in the conflicting flow. Also known as a permissive movement (MUTCD term). *See also Movement.*

Phase—A timing unit associated with the control of one or more movements. Phases are often assigned to vehicular and pedestrian movements.

Phase Failure—The occurrence of one or more stopped vehicles that cannot proceed through a signalized intersection on a single green indication.

Phase Pair—A combination of two phases allowed within the same ring and between the same barriers (i.e., 1+2, 5+6, 3+4, and 7+8).

Phase Sequence—The order of phases in a ring.

Phase Table—The timing parameters associated with a phase.

Phasing Diagram—A graphical representation of a sequence of phases, typically in the form of a ring-and-barrier diagram.

Platoon—A group of vehicles or pedestrians traveling together as a group, either voluntarily or involuntarily because of signal control, geometrics, or other factors.

Power Supply—An electrical device that converts AC to correct DC voltage for various devices in the signal cabinet.

Practitioner—A general term for anyone responsible for signal timing.

Presence Mode—A detection mode where a signal is sent to the controller for the duration of time a vehicle is inside a detection zone. *See also Pulse Mode.*

Pretimed Control—A mode of operation where every phase is on recall every cycle, regardless of changes in traffic conditions. Typically only used in closely spaced grid-like networks.

Protected Movement—A movement that has the right-of-way, and there are no conflicting movements occurring. *See also Movement.*

Protected-Permitted Left-Turn Phasing—Compound left-turn protection that displays the permitted phase before or after the protected phase.

Pulse Mode—A detection mode where vehicle detection is represented by a single brief "on" pulse to the controller. Typically only used for counting traffic, as presence detection protects against premature gap out due to vehicles queued over the detection area. *See also Presence Mode.*

Queue—A line of motorized vehicles, bicycles, or pedestrians waiting to be served by the system (i.e., stopped) at the beginning of the green/walk interval. Slowly moving vehicles or pedestrians joining the rear of the queue are usually considered part of the queue.

Queue Discharge—A flow in which queued vehicles start to disperse (with high density and low speed).

Queue Spillback—A term used to describe the vehicles stopped at an intersection that exceed the available storage space for a particular movement.

Recall (or Phase Recall)—A call is placed for a specified phase each time the controller is serving a conflicting phase. This ensures that the specified phase will be served again. Types of recalls include maximum, minimum, pedestrian, and soft. *See also Maximum Recall, Minimum Recall, Pedestrian Recall, and Soft Recall.*

Red Clearance Interval—The period of time following a yellow change interval, indicating the end of a phase and allowing additional time before the beginning of conflicting traffic.

Red Rest—All phases may rest in red state when no serviceable calls exist in the ring and no other recall modes are used.

Red Time—The period of time (expressed in seconds) in the cycle during which the indication for a given phase is red.

Ring—A sequence structure consisting of two or more sequentially timed and individually selected conflicting movements, arranged to allow flexibility between compatible movements in different rings.

Saturation Flow Rate—The equivalent hourly rate at which vehicles can traverse an intersection approach under prevailing conditions, assuming a constant green indication at all times and no lost time (in vehicles per hour or vehicles per hour per lane).

Saturation Headway—The average headway between vehicles occurring after the fourth vehicle in the queue and continuing until the last vehicle in the initial queue clears the intersection.

Semi-Actuated Control—A type of signal control where detection is provided for the minor movements only, and the signal timing returns to the major movement because it has no detection and is placed in recall. This is typical for coordinated operations without coordinated phase detection. *See also Actuated Signal Control.*

Setback Detection—Detection located upstream of the stop bar. Allows for more efficient phase termination and is used for decision zone protection. If stop-bar detection is not present, setback detection will use locking memory mode. *See also Detector.*

Signal Monitor—A safety device that monitors signal outputs for conflicting or improper outputs. Types include a simple conflict monitor; however, modern signal monitors perform many additional functions beyond looking for conflicts. Other types include malfunction management units (MMUs) and conflict monitoring units (CMUs), possibly in combination with auxiliary monitor units (AMUs).

Simultaneous Gap Out—This parameter requires all currently green phases to concurrently “gap out” prior to crossing the barrier. Generally only recommended for high-speed rural-intersection through phases. *See also Gap Out.*

Soft Recall—The recall parameter that causes the controller to place a call for vehicle service on a phase in the absence of a serviceable conflicting call. *See also Recall.*

Software-in-the-Loop (SITL)—A means of providing a direct linkage between simulation models and software emulations of controllers.

Speed—A rate of motion expressed as distance per unit of time.

Split—The time assigned to a phase (green and the greater of the yellow change plus red clearance or the pedestrian walk plus clearance times) during coordinated operations. May be expressed in seconds or as a percentage.

Split Failure—*See Phase Failure.*

Start-Up Lost Time—The additional time (in seconds) consumed by the first few vehicles in a queue at a signalized intersection above and beyond the saturation headway, because of the need to react to the initiation of the green phase and to accelerate. *See also Lost Time.*

Steady Don’t Walk—The period of time after the walk and flashing don’t walk have completed timing.

Stop-Bar Detection—Detection located at the stop bar that is generally used to discharge the queue. Generally uses non-locking memory mode in combination with large area detection design, so that vehicles making permitted turns do not extend or call the phase. *See also Detector.*

Switch Pack—*See Load Switch.*

Time before Reduction—This volume-density timing period begins when the phase is green and there is a serviceable call on a conflicting phase. When this period is completed, the linear reduction of the passage time begins until the minimum gap is reached or the phase terminates due to gap out. *See also Time to Reduce.*

Time-of-Day Plans—Signal timing plans associated with specific hours of the day (i.e., associated with fluctuations in demand), days of the week, or days during the year (e.g., holidays, seasons).

Time-Space Diagram—A chart that plots the location of signalized intersections along the vertical axis and signal timing along the horizontal axis. This is a visual tool that illustrates coordination relationships between intersections.

Time to Reduce—This volume-density timing period begins when time before reduction ends and controls the linear rate of reduction until the minimum gap is achieved. *See also Time before Reduction.*

Traffic Management Center (TMC)—An optional physical component of a signal system, which contains the operational database that stores controller data, allows monitoring of the system, and allows timing and other parameters to be modified.

Trailing Overlap—A signal indication that ends after the parent phase. An example application of a trailing overlap is providing interior clearance at two “offset-T” intersections. *See also Overlap.*

Travel Time (Average)—The total elapsed time spent traversing a specified distance. The average travel time represents an average of the runs for a particular link or corridor.

Uncoordinated (or Free Operation)—A traffic signal not operating as part of a coordinated system of intersections. Free operation can be set by time of day.

Uninterruptible Power Supply (UPS)—A backup battery power supply to operate a traffic signal during power outages

Unit Extension—*See Passage Time.*

Upstream—The direction from which traffic is flowing.

User—A person (pedestrian) or specific type of vehicle (e.g., bike, transit, or truck) that uses a traffic signal.

User Priority—A user may be assigned a relative or absolute priority based on operating environment and locally desired outcomes. These priorities may vary by movement.

Variable Initial—An interval that times concurrently with the minimum green interval and increases by each vehicle actuation received during the initial period. This time cannot exceed the maximum variable initial.

Vehicle—Every device in, upon, or by which any person or property is or may be transported or drawn upon a highway, excepting devices used exclusively upon stationary rails or track (*Uniform Vehicle Code*, 2000).

Volume—The number of pedestrians or vehicles passing a point on a lane, roadway, or other trafficway during some time interval (often 1 hour), expressed in vehicles, bicycles, or pedestrians per hour.

Volume-Density—A phase timing function that uses parameters (e.g., variable initial, minimum gap, time before reduction, and time to reduce) to provide appropriate minimum green time to clear intersection queues when stop-bar detectors are not used, and/or it is desired to adjust the passage time. *See also Density.*

Volume-to-Capacity Ratio (or Degree of Saturation)—A ratio of demand volume to the capacity for a subject movement.

Walk Interval—An indication providing initial right-of-way to pedestrians during a pedestrian phase and prior to the pedestrian clearance interval.

Yellow Change Interval—An indication warning users that the green, flashing yellow, or flashing red indication has ended, and the red indication will begin.

Yellow Trap—A condition that leads the left-turning user into the intersection believing the opposing user is seeing a yellow.

Yield Point—The earliest point in a coordinated signal operation that the controller can decide to terminate the coordinated phase(s). It is typically followed by one or more permissive periods that allow the controller to yield to uncoordinated phases later in the cycle, yet still return to the coordinated phase(s) in time to remain in coordination. Permissives are primarily beneficial during lower traffic volumes, allowing uncoordinated phases to be served if they arrive later than the initial yield point.

Abbreviations and acronyms used without definitions in TRB publications:

A4A	Airlines for America
AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
HMCRP	Hazardous Materials Cooperative Research Program
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
MAP-21	Moving Ahead for Progress in the 21st Century Act (2012)
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
PHMSA	Pipeline and Hazardous Materials Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S.DOT	United States Department of Transportation

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001

ADDRESS SERVICE REQUESTED

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.
www.national-academies.org

ISBN 978-0-309-30888-5
90000
9 780309 308885

NON-PROFIT ORG.
U.S. POSTAGE
PAID
COLUMBIA, MD
PERMIT NO. 88