The
"WHY"
and
"HOW"
of Automatic Train Operation

By Franklin George
Assistant to Executive Vice President
General Railway Signal Company

Reprinted from the June, 1961 issue of Metropolitan Transportation
A Hitchcock Publication
The “WHY” and “HOW” of Automatic Train Operation

By Franklin George
Assistant to Executive Vice President
General Railway Signal Company

Every time the subject of automatic train operation comes up, there usually arises two questions which, in their logical order, are “Why?” and “How?” Of course, the third question, if not sometimes the first question, is “How much?”, but I won’t attempt to go into that at this time.

We don’t feel that anyone is going to buy ATO just for automation itself, but rather for the benefits that will accrue from its use. For rapid transit rail lines I suggest the following potential benefits as to the “why” of ATO:

1. Properly applied and integrated with other efficient and attractive facilities and equipment, it will enable the transit system to provide the kind of service that will attract, and retain, the patronage of a greater percent of the traveling public.

2. It will enable the transit systems fixed plant and rolling stock to be employed in gainful service more hours of the day than is now economically feasible.

3. If properly organized system wide, it will enable better and more simply accomplished supervisory control of train operations—there will be less people required at fewer points to observe and make decisions regarding operation.

4. It will insure more consistent “on time” operation.

5. It will minimize dependence on availability of qualified train operating personnel—and on the human element as regards safety of operation.

6. It will reduce cost per passenger mile of service rendered.

No degree of automation can attract and hold patronage if the train system lines do not run between the points the public desires to travel between, if the stations are inconvenient to get to, or, in the suburbs, if there are not ample automobile parking facilities. Also, the stations and the cars themselves must be reasonably comfortable and attractive, with ease of entry and exit, even with an armful of bundles. And, equipment and plant must be such that for the longer hauls an average speed which is better than that possible with the private automobile must be obtainable. In brief, I am saying that you can’t just take an obsolete, run-down transit line and its equipment and make it pay with automation alone.

Will Make Off-Peak Service Possible

I believe that with ATO it will be economically possible to provide a frequency of service during the off-peak hours, say from 9:30 a.m. thru 3:30 p.m. and from 7:00 p.m. until 1:00 a.m. or 2:00 a.m., that cannot be otherwise provided except at high cost and out-of-pocket loss.

Such service, if it serves the right areas and is dependable, may very well attract the local businessman, the housewives and shoppers, and the evening “downtowners” and “suburban visitors.” Today, many
of these people either don't go, or they drive their cars, mainly to be able to go and come at times that suit their convenience—not because they enjoy driving in city traffic.

A rapid transit system's fixed plant, track, stations and parking lots, exist—waiting to be used—24 hours a day, 365 days a year. So do the trains. And, idle plant and trains do not contribute one penny toward reducing the debt that was incurred in obtaining them, the interest thereon, nor the fixed maintenance charges which go right on accumulating every minute of every day, whether the facilities are busy or idle.

If ATO is applied to an entire system, or a large segment thereof, I anticipate there would be desire for a centrally located supervisory train operation monitoring and control facility. I do not mean by this the provision of means to observe precisely where every train is, and what exact speed it is traveling at a specific moment, nor the means to remotely control the individual train's rate of acceleration, deceleration, or precise speed, but rather the means to observe how the over-all system is functioning generally. And, if any abnormality of consequence occurs, be able to reach in and do something about it—stop additional trains from entering the troubled area, reroute affected traffic, put on extra trains, or initiate such other pertinent action as may be required. Such a supervisory center should be close to source of information concerning traffic demands and the determination of how to satisfy those demands. It should be known, for example, that the ball game should be over about 5:30 p.m. and there will be approximately 6,000 people desiring to come back into town at that time.

With a properly organized system, the local portions thereof—out where the conditions exist and the specific detailed action is required—will automatically and locally take care of controlling train speeds, stops, and starts so as to avoid creation of unsafe conditions and yet function to obtain the benefit of all operating advantages possible consistent with best possible on-time performance.

I believe that once the public becomes educated to the fact that the trains' station leaving times are automatically programmed, they won't be lackadaisical about strolling down the platform as they are when they know the conductor won't wave the high-ball until he sees them on the train. They will respect the trains starting time, just as people do after they've lived with automatic elevators for a while.

Normal Train Operation Programmed

If normal train operation is programmed and automated properly, the dependence on availability of skilled train operating personnel will be lessened. Operating records indicate that a high percent of accidents, minor and major, can be attributed to man failure. That is why on many railroads, signals, train
stops and even cab signals and speed control, have been installed—to safeguard, continuously and consistently, against man's occasional fallibility.

If the trains themselves operate automatically, labor costs will be reduced there. If the kind of service rendered attracts and holds more riders, there will be more passengers carried every car mile. Furthermore, if the ATO is organized on a system basis, there will be fewer manned interlockers and the like along the right of way. A smaller operating supervisory staff will be able to do a better job. All of these will contribute to lower operating costs.

So much for the "why" of considering ATO in our concept of future rapid transit operation. As what we now call the future becomes reality, the benefits of experience will no doubt make the picture clearer so that we will better recognize ATO's potential benefits and see the economic reasons why we should automate such operations as can be automated.

The "How" Of ATO

Now, "how" do we apply automatic train operation to rapid transit operation. The first thing that comes to our mind is, as the name implies, the fact that we are going to operate the train automatically, or do automatically those things a motorman does today. We should, therefore, analyze what a motorman does do. Basically, he exercises his five senses—sight, hearing, smell, taste and touch or "feel." And then, by applying good judgement, based on his knowledge of conditions and his experience, he performs such action as he deems necessary and desirable at the time.

With sight, he observes his location, and conditions affecting operation of his train which are visually apparent. Sometimes there are severe restrictions on what he can and should see, so a long time ago we supplemented this sense by providing signals and signal systems which automatically and continuously observe conditions within and beyond man's ability to perceive. Sight is also used to sense speed.

With hearing, he detects and identifies familiar and unfamiliar sounds, keeping alert to detect if any unfamiliar sound may be indicative of some abnormality which could affect the safety of operation. In some instances, he may receive operating instructions or other pertinent information by means of his hearing. Hearing also aids the sensing of speed, particularly when visibility is limited.

With smell, he detects and identifies familiar and unfamiliar odors, again keeping alert to detect any that might be indicative of trouble—like the smell of burning insulation, for example.

Taste may not be of quite so much consequence, although I assume it does to some degree supplement the sense of smell.

The sense of touch contributes to the evaluation of rates of acceleration, deceleration, and actual speed. Track and car condition abnormalities are many times first sensed by the "feel" of the ride.

Technology Can Duplicate Senses

By employing the technological capabilities available to us today, we can provide practical means for acquiring all of the facts pertinent to train operation that I referred to—in fact, some that can't be sensed by man alone. And by employing decision making equipment, which in turn bases the decisions on predetermined logic applied to the facts, we can initiate action of the kind, and in the sequence, best suited to the individual situations as they arise. Once we have attained this plateau of ability, all that remains is the simple mechanization of the required operating procedures.

How do we do these things; sense the necessary facts, make the desired decisions, and accomplish the required operating procedures?

We hear, and read, of many schemes for automating operation of trains. I shouldn't really say "schemes," because many boil down to simply assertions that it can be done—by just putting a black box on the train to take the place of the motorman, provide it with a radio and radar whose transmissions can go thru tunnels, around curves, and thread themselves thru a maze of metal structural members, and you are in business. Well, it isn't quite that simple.

The logic I suggest applying to the problem is: first, determine what you have to know, how you sense or obtain this data, and then what you do as a result of it.

First, we have to have a means of monitoring wayside conditions affecting train operation. Some of the more important of these conditions are the continuity of the track and of the rails themselves, track occupancy or presence of trains or cars, switch positions and route alignments and/or availability. And, in some areas, it may be desirable to monitor or check for smoke, fire, or water. There might even be need for detecting movement or rock, as in tunnels or rock cuts, or of earth movement or shifts which could affect track alignment or foul movement thereof. In station areas, it might be desirable to check for foreign objects on the track, such as
trespassers. It may also be necessary or know the identity of approaching trains, and how many cars they may have.

If we are talking about a rail transit system, then the track circuit can provide us a fail-safe means for monitoring rail continuity. If we use steel wheels on steel rails, then the track circuit automatically provides us a means for detecting track occupancy, thereby warning us of the presence of trains or cars which would foul or obstruct the movement of other trains or cars.

If we don’t have trains with steel wheels on steel rails, then we must substitute for the track circuit other means for detecting “block” occupancy, or presence of trains or cars. A system, organized on the fail-safe principle which detects the passage of a car or train past fixed wayside points without the necessity of mechanical or electrical contact with the cars or trains themselves, is available. It can detect that a train has entered a section of track or zone, and it can detect when the train leaves, that it completely leaves, that zone. By coordinating the indication of entry into one zone with the exit of the preceding zone, a closed-circuit fail-safe scheme of track or zone occupancy is obtained.

The proper position and safe condition of track switches are continuously monitored by controllers mechanically connected to hand or spring-operated switches, and by the switch throwing and indicating mechanism on power-operated switches. By integrating this information, locally, with information concerning other track conditions, presence, or absence, of cars or trains, either in the immediate area, passed or approaching, we obtain information as to route alignment and availability. By utilizing information as to identity of approaching trains, or of their individual destinations, we establish the routes desired for them and, if our system is sufficiently sophisticated, we establish the normally preferred routes. If these are not available, the next best alternate routes and so on, if any are available, can be established. If none are available, we store the “route desired” information until desired routes become available, and then set them up. These decisions and actions are performed automatically with predetermined logic inherent in present-day signaling and automatic interlockings.

Now that we have sensed locally all wayside factors governing train movement, we integrate the resulting data with any required central intelligence commands or requisites, and then locally make the decisions as to the manner of commands or orders we need to give to the train or trains approaching or operating in the area of the local control. For example, if we have a train approaching for which there is no proper route available because of more immediate pending passing of another train whose movement would conflict with that of our approaching train, then our decision must be that we tell the approaching train to stop short of where it would conflict with the passing train. This, in turn, initiates automatically telling the approaching train to start slowing down farther out on the approach in preparation for making the required stop.

Now that we have made our decision as to what we should tell the approaching train, how do we convey the message to it?

First, we have to convert the message, or command, into a form of intelligence, or code as we like to call it, that can be handled by our wayside-to-train communication system and that can be interpreted on the train into specific commands to be acted upon by the train equipment. The codes may be comprised of simple time periods of energy, “on” and “offs,” with the lengths of time periods varied for the different messages; or they may be comprised of modulations applied to a carrier frequency, also time coded and used in combinations, whichever best suits a particular installation.

Specific Messages for Specific Trains

If we have the steel wheel on steel rail type of trains, we can use the same medium for sending messages to the train as we use for detecting the presence of trains—the track circuit. If other, such as rubber tired wheels on concrete type of track, then we substitute a continuous wire for the rails.

In either case, the system philosophy used is that messages should not be broadcast, but that a specific message should be discretely delivered to a specific train while it is in a specific section of track, or block. In other words, it should be that message to that train calling for that action to be taken at that place at that time. And the transmission and reception thereof on the train must be continuous. Loss of reception of any proceed authorization message automatically forces a train stop—in fact an emergency stop—if a message calling for a service brake stop is not substituted for a proceed message.

Where we employ the track circuit, we send the message via energy in the rails, one rail forming the outgoing line and other the return line. The message is inserted at the end of the block toward which the train is moving, and it is not inserted into a block until a train has entered that block. As the train moves on into the block, it shuts out the message thereby preventing it from reaching any other train to the rear. The energy, or carrier, used in the rails must be of character, frequency-wise, that is as far removed as possible from existing power frequencies or harmonics thereof, or any other suspected foreign current. Where we have a track for each direction, a different frequency may be employed for each
direction. The energy level used should be that which will reliably override the "noise" from the traction currents which are in the rails.

The foregoing concept, or organization, of the wayside-to-train communication system pretty well safeguard against extraneous influences and incorporates the closed-circuit, fail-safe principle of operation. The same concept is, in essence, followed where we cannot have track circuits and must employ a paralleling "loop" wire instead of rails.

The train equipment consists of apparatus for receiving and decoding the messages, for monitoring and determining actual train speed and, using information from these two sources, for formulating the manner of action the train's power and brake equipment should perform. Such action must also be governed by train equipment sensing devices, that is, for example, if air brakes are used, then there must exist a minimum brake system air pressure or the train will be stopped or, if stopped, prevented from starting. Also, doors must be closed, power must be on, and so on.

The train's receiving apparatus consists of coils having laminated sheet steel cores, a filter and an amplifier. One receiver coil is used over each rail, or one over a loop wire if not a track circuit installation. The equipment is sensitive enough to pick up its desired messages with several inches separation between the rails and the loop wires, but not sensitive to the point where messages from other tracks or extraneous sources would be picked up. With this degree of sensitivity, and this amount of physical separation, car sway or bounce will not adversely affect operation.

Only those messages received via the selected carrier frequency are amplified. Following the amplification, there is the decoding equipment with its tuned filters and associated decoding circuits, all arranged to provide selective response only to the assigned code rate, or rates, as the case may be. If carrier modulation is used, the code would be comprised of combinations of two or more modulation frequencies, coded or sent at the assigned rate. The modulation frequencies would be varied in accordance with the commands being transmitted. The wayside commands so received, amplified, and decoded, cause corresponding relays to be energized.

Working in conjunction with this part of the system, we have the speed detecting apparatus. This consists of a car axle driven frequency generator, an oscillator, an amplifier and a number of high pass filters and relays, depending on the number of speed ranges we desire to identify and mark.

Following our closed-circuit, fail-safe principle of operation, we require one more element in our system of train control—the means to monitor or detect on the wayside the train's performance. Is it obeying the commands, or is it getting out of control because of some abnormality?

Such monitoring, or checking, can be done in several ways, with various degrees of authenticity. Precise speed checks may be made when the train passes fixed wayside points, but this requires interpretation and repeat checks to be of much value. The train can be equipped to transmit to the wayside a "condition report," a report which can tell the wayside what the train speed is and if the major items of control equipment are functioning properly. This train to wayside message can be picked up intermittently, at selected fixed wayside points or, in a more sophisticated arrangement, be picked up continuously on the wayside. The latter holds more promise of being a full closed-circuit, fail-safe scheme. On the wayside, the receipt of an abnormality indication, or rather the non-receipt of a normal conformation report, would automatically initiate there such action as required to stop or otherwise safeguard the train in trouble.

Centralized location supervisory control can be elaborated and sophisticated as desired, or, more realistically, as the pocketbook will permit. But, I suggest caution in this regard. If a small system or segment of a system is being automated, it may be economically feasible to centralize more of these details, and a small supervisory staff, one man for example, may well be able to provide all of the required supervisory control. In a large system, however, I recommend localizing all possible condition monitoring and the decision making and command forming equipment to cut down the dependence on extended and complex communication facilities. Give the supervisory control center only a condensed but pertinent over-all operating picture—that which is essential to maintaining orderly traffic flow.

Should Also Have Two-Way Voice System

I would also recommend that the system, and the trains themselves, be provided a good two-way voice communication system. I assume we would not want completely unmanned trains, for social reasons more than technological. An employee on the train could provide the supervisory control center any manner of situation report. He should also be provided means for initiating an automatic train stop. I would also recommend that he should be provided a normally sealed means of moving the train at a low speed, not to exceed five miles per hour, when such movement is properly authorized by the control center. In this way, you would preclude possibility of stranding passengers in a tunnel or between stations.

Automation will enable transit systems to provide more attractive service, at less cost per passenger mile.