

Back to basics: Interlocking

Part 1



Francis How



This, the third in a series of articles on 'back to basics' themes, looks at the essentials of 'interlocking'. Interlocking is central to railway signalling, as it ensures that the components of a signalling system act together in a manner which is safe for the routing and movement of trains. Whole books could be written about the subject, and this article is no more than an introduction, intended for IRSE members new to the industry rather than those who are experienced in specifying, designing and testing signalling systems.

This article focuses primarily on the technology used for interlocking. Next month there will be a further article, in which we will look at the functionality of an interlocking – what it actually does in practice.

What is interlocking?

If you ask a signal engineer about interlocking, they may well point to an equipment room full of relays, or a cabinet of computer equipment in a control centre, or perhaps even some complicated-looking arrangement of metalwork underneath a lever frame in a signal box. It is true that all these things are 'interlockings', but 'interlocking' is defined as a feature of a control system that makes the state of two functions mutually dependent. In the context of railway signalling, interlocking is used to keep trains safe from collision and derailment. The primary purposes of these interlocking features are to ensure that:

1. Before a train is given authority to move along a section of track from one signal to the next:
 - points are in the correct position (to avoid derailment),
 - there are no trains already on the track (to avoid collision), and
 - no conflicting train movements are already authorised (also to avoid collision).

"Interlocking does not check that everything is safe for the passage of a train"

2. When a train has been given authority to move:
 - points in the section of track are prevented from being moved, and
 - other trains are prevented from entering the same section of track.

until the train has passed through the section of route.

There is more to interlocking than this, as we shall see, but this is the essence of what it is all about.

What does interlocking not do?

Interlocking does not check that **everything** is safe for the passage of a train. A section of railway track must be safe for the passage of trains in many other ways as well. For instance, the distance between the rails must be correct, the track-bed must be capable of supporting the weight of a train, and the train's cross-sectional dimensions (the dynamic envelope) must be compatible with the positioning of items such as platforms and bridges, so that the train will not hit them. These can also cause collision or derailment – but they are not generally the concern of the signal engineer. Railway engineers have other methods for ensuring the integrity of these other features upon which train safety depends.

Interlocking safety

Interlocking functions (such as moving a set of points or clearing a signal) must be executed only when it is safe to do so. Industrial control systems (of which railway signalling is an example) are designed to meet a specified 'safety integrity level' (SIL). There are five such levels, from 0 to 4, and the interlocking functions in a modern main line or metro railway signalling system must usually meet SIL 4 requirements – the highest level possible. This means the likelihood of an unsafe failure is incredibly small.

The underlying safety principle traditionally associated with railway signalling, and particularly with the interlocking, is known as the 'fail-safe'

principle. This means that if an interlocking system develops a fault, it is designed so that it will fail in a manner which stops trains, by putting signals to danger. This fail-safe property is achieved in various ways, including the use of inherently fail-safe components, the design of the interlocking logic, and the system architecture.

It is important to note that 'fail-safe' does not mean that the signalling of trains is 100% safe. This is partly because in practice the occurrence of unsafe failures cannot be completely eliminated, and partly because if trains have been stopped by a (safe) failure of the signalling system, the movement of trains then depends largely upon the application of operational procedures, with the associated risk of human error.

Not all parts of a signalling system need to be ultra-safe. In modern systems the use of high integrity (SIL 4) design techniques is usually restricted to those parts of the system which are essential for safety – including the interlocking functionality. Other parts, such as the control panel or desk, are usually of a lower integrity (typically SIL 0 to SIL 2). You might wonder why we do not design all parts of the signalling system to achieve SIL 4 levels of safety. The answer is that designing systems to meet high levels of safety integrity is complex, time-consuming and expensive, and can lead to lower levels of reliability.

A little history...

When railways first appeared, they had no signalling in the form that we would recognise today. The concept of a signal box did not exist, signals were very rudimentary (originally just a man with a flag), and giving permission for a train to enter a section of track relied simply upon allowing sufficient time for the preceding train to

Figure 1 – The Chippenham, UK, factory of Westinghouse Brake & Saxby Signal Company in 1927. Some 60 years after the invention of the interlocking skilled teams were still assembling complex lever frames – some of which are still in use today.

*Photo WB&S Archive/
Chippenham Museum &
Heritage Centre.*

have left the section – without any knowledge of whether in practice it had done so!

Not surprisingly, it didn't take long for accidents to demonstrate the need for safer ways of controlling the movement of trains. This article does not explore the evolution of railway signalling, but there were some key milestones in its development which are worth noting. One was the introduction of the electric telegraph, so that someone at one end of section of railway line could communicate with someone at the other end. This eliminated the need for 'time interval working' by enabling the person controlling entry to a section of line to receive positive confirmation when the whole of the preceding train had left the section and therefore the line was again 'clear'.

A second crucial development was the introduction of the signal box, to enable signals and points to be controlled from one place. Previously the setting up of a route for a train had relied upon people walking about on the track to move points and operate signals. Centralisation of these activities was not only more efficient and reduced the possibility of misunderstanding, it also facilitated the introduction of 'interlocking' – mechanical equipment in the signal box that helped to prevent mistakes being made by the signaller when moving points or operating signals. Even in the age of computerisation, it is remarkable to look back at how railway engineers of the 19th century invented mechanical logic systems that largely overcame the risk of human error when signalling trains.

So, the way was paved for the introduction of 'interlocking' according to a defined set of principles or rules – which for the most part still apply today, albeit they vary in some respects from railway to railway.



Interlocking technologies

Mechanical interlocking

The earliest form of interlocking was purely mechanical. Signal boxes were usually two storey buildings, with the signaller working upstairs and the interlocking downstairs. To allow a train to move, the signaller would operate large levers, separate ones being provided for each set of points and for each signal. These levers were directly connected via mechanical rods and/or wires to the points and signals outside, and therefore could require considerable effort to move.

On the ground floor of the signal box, underneath the levers, was an arrangement of metal bars that were connected to the levers, with other bars at right-angles to the first. The physical interlocking of the two sets of bars prevented the signaller from moving a lever unless other levers were in the correct position. So, for instance, he could not move a signal lever to allow a train to move unless the relevant points levers were in the correct position.

Gradually engineers began to introduce simple electro-mechanical locking in combination with the mechanical locking, to prevent a lever from being moved unless its electric lock was energised. Energisation of the lock (via an electrical circuit) required other levers to be electrically proved to be in the correct position. Further conditions were progressively added to the lock circuits to improve safety, such as the requirement for relevant track circuits to be clear before a signal or set of points could be moved.

Safety features were also added to the block systems that controlled the movement of trains between neighbouring signal boxes, to prevent a signaller from sending a train from the area that he controlled to that of the next signal box unless it was safe to do so.

Mechanically interlocked signal boxes are still in use on railways in many parts of the world, well over 125 years since they were introduced.

How does mechanical interlocking work?

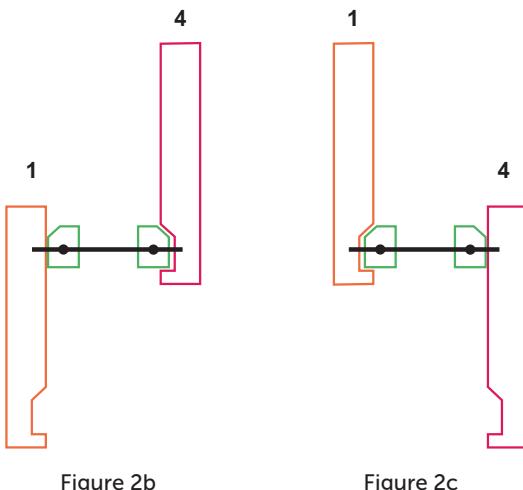
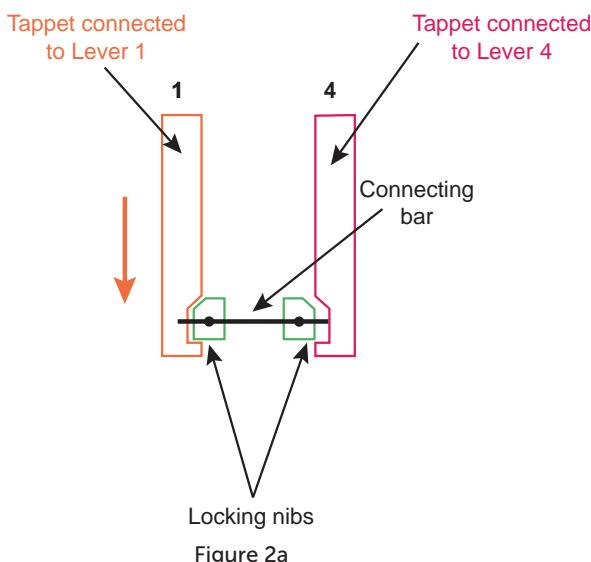
A flat metal bar (called a tappet) is attached to the end (tail) of each lever. All the tappets are held within a locking box, so that each one moves in one direction when the corresponding lever is pulled to its reverse position, and in the opposite direction when the lever is pushed back to its normal position.

To create a lock, a bevelled notch is cut in the side of a tappet, and a locking piece (sometimes called a nib) is cut to fit the notch. If a horizontal bar is placed in the locking box with one end connected to the locking piece, and the other end is connected to another locking piece, the movement

of one lever is prevented or permitted according to whether a locking piece is held in a notch in the tappet, or is clear of (or free to move out of) the notch. The use of bevelled edges enables a tappet, when free to move, to force the locking piece out of the notch.

In Figure 2a below, when Lever 1 is pulled from normal to reverse, the tappet will move in the direction of the arrow. The positions of the tappets and locking pieces will then be as shown in Figure 2b, and Lever 4 is locked in the normal position. It cannot be moved because the second locking piece (connected to the first) has engaged in

the notch in the tappet of Lever 4 and is not free to move out of it. Thus Lever 1 locks Lever 4. The converse is also true. If Lever 1 is normal and Lever 4 is reverse, Lever 4 locks Lever 1 normal, as shown in Figure 2c (this is called reciprocal locking and is an inherent feature of mechanical locking). Much more complicated locking arrangements can be created than the simple one shown, with levers being locked in both normal and reverse positions by multiple other levers, and with locking of one lever by another being conditional on the position of a third lever.



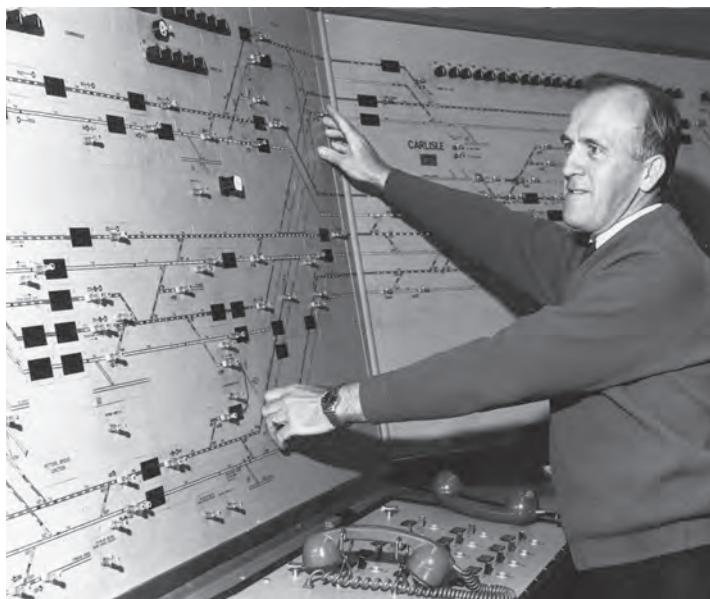


Figure 3 – The introduction of relay-based signalling allowed control centres to move to the use of complex panels. These UK examples are from Carlisle (left) and London Bridge (right) power signal boxes, commissioned in the 1970s.

Photos WB&S Archive/
Chippenham Museum &
Heritage Centre.

Figure 4 – Different approaches to relay design. Below, the type N relay, bottom the type C. Photos Siemens Mobility
and E Dold & Söhne KG.



Electrical interlocking

The advantages of using electrical locking became even more evident when colour light signals and electric point machines began to be introduced. Mechanical locking could be entirely replaced by electric locking, and large levers were no longer necessary for operating points and signals, since no great physical effort was required by the signaller to switch an electric signal or point machine.

Early all-electric signal boxes used miniature mechanical levers on a desk, replicating in a more modern form the row of levers in mechanical boxes. This evolution led in time to the introduction of control panels, with switches and buttons on the panel being used to set whole routes from one signal to the next, without the signaller needing to set the points in the route individually.

The 'route setting' approach eventually became the preferred form of control. The role of the interlocking was crucial in this. Instead of being a passive system for determining whether it was safe to operate a set of points or a signal (as mechanical locking had been), it became an active system that interpreted and acted on requests received from the control panel. In simple form, the request to set a route from one signal to the next is set up by the signaller using switches and buttons on the control panel (which has a diagram of the track layout on it). The interlocking then moves the relevant points provided it is safe to do so, checks that the track is clear of other trains, and clears the entrance signal for the route.

Relays

The fundamental building block of the traditional route-setting interlocking is the relay. Before the relay interlocking gained prominence, relays had already been used for track circuits and for other simple circuits in mechanical and early electric signal boxes. These relays had generally been relatively large devices, often designed to sit on shelves. But as relay interlockings became more popular, relays were progressively made physically

smaller, and were often designed to plug into bases to which all the wiring was connected, so making it easy to replace a faulty one as well as enabling hundreds or even thousands of them to be housed in a much smaller space. All the wiring and the relays and their bases were mounted on vertical racks, and a large interlocking might have dozens of such racks, all housed in a 'relay room'.

Not sure what a relay is?

A relay is an electromechanical switch, with an electromagnetic coil, an armature and various contacts. When the coil is energised, the armature moves and closes a number of contacts ('front' or 'normally open' contacts) and opens others ('back' or 'normally closed' contacts). These contacts are used in other circuits to create the logic conditions for operating other relays, powering point machines, illuminating the aspects of signals etc. When the coil is de-energised the armature returns to its original position, opening the front contacts and closing the back contacts.

Relay interlocking architecture and design

The architecture of relay interlockings varies from railway to railway (even within a single country), and from country to country. We are not going to explore all the variations here, but it is worth understanding a little about the basic design philosophies that characterise almost all relay interlockings.

In the early days all relay interlockings were 'free wired'. With this approach, each circuit, whatever its purpose or function, was individually designed and wired, usually in accordance with a set of templated (standard) circuits. In time an alternative approach emerged, whereby manufacturers provided a range of factory-wired, pre-tested sets of relays known as 'geographical' units. Each type of unit was designed to provide the standard interlocking functions required for a specific combination of signals, points and train detection sections. By connecting the appropriate units together (usually with plug-coupled cables) to mimic the actual layout of the track and signalling, the required route-setting functionality could be

Figure 5 – Route relay interlockings have been very successful in a huge range of applications world-wide. In this example from the original Singapore MRT scheme the interlocking is interfaced to coded track circuits allowing a high performance automatic train protection and automatic train operation system to be implemented.

*Photo WB&S Archive/
Chippingham Museum &
Heritage Centre.*



built up relatively easily. There are advantages and disadvantages with both approaches, including cost, flexibility, speed of design and testing.

Secondly, the relay interlocking circuits (both free-wired and geographical types) vary in form according to the type of relay used. There are, broadly speaking, two generic types of relay used for all interlockings. One type is inherently 'fail-safe', meaning that if the coil is de-energised, the front contacts will always open, and it is virtually impossible for a failure to occur whereby front and back contacts are in the 'closed' position at the same time. The use of non-welding materials for the contacts makes it impossible for a contact to weld in the closed position. This type of relay is known generically as type N in UIC standard 736i. The best-known family of signalling relays in this category is probably the BR930 series, the development of which, incidentally, involved the IRSE. There are at least 200 variations of this same basic relay, with different operating characteristics (slow to energise, slow to de-energise etc), different numbers and types of contacts, and different operating voltages.

The other generic type of relay used in some interlocking systems is known as type C in UIC standard 736i. It is not guaranteed to behave in the inherently fail-safe manner described above. Specifically, it is possible for a contact to weld so that it remains closed when it should be open although, as with a type N relay, the mechanical design prevents any front and back contacts being closed at the same time, even if welding occurs (a feature known as 'forcibly guided' contacts). Such relays have the advantage of being considerably less expensive and smaller. But in order for the interlocking as a whole to behave in a fail-safe manner, the circuits are more complicated as a consequence of using additional contacts to prove that relays have de-energised correctly, and because of the need to check that the circuits are operating in the correct sequence

with the passage of the trains. By contrast, the dependable fail-safe nature of the type N relay makes it generally unnecessary to include this additional complexity.

In all interlockings the circuits are designed to exploit the safety characteristics of the relays. Usually this is done by requiring a relay to be energised to allow a less restrictive state (e.g. to allow a signal to show a proceed aspect, or to allow a set of points to move). If the relay or the power supply fails, or there is a disconnection in the circuit, the relay de-energises, so causing the signalling equipment to revert to a safer state.

Computer-based interlockings (CBI)

With the development of electronic logic gates in the form of integrated circuits, and subsequently with the emergence of the microprocessor and programmable logic controllers (PLCs), it was a natural step to see how this technology could be applied to interlockings. Early experimental installations were implemented in the 1960s and 1970s, but it was in the mid-1980s that the electronic software-based interlocking first became a reality. One of the best known of these was SSI – the 'Solid State Interlocking', developed in the UK.

The use of software-driven electronic logic presented a whole new set of challenges for system designers. A route relay interlocking is, in effect, a hard-wired parallel processing logic machine. If it goes wrong it could initiate unsafe actions, but the potential failure modes and their causes are well-understood and, by good design practice and by testing it to make sure the locking conforms to the application rules, the probability of an unsafe failure is very low. A computer-based interlocking (CBI), which makes use of microprocessors, is another matter entirely, however. Microprocessors comprise hardware and embedded software, and these are not designed to meet the high integrity safety requirements

"A route relay interlocking is, in effect, a hard-wired parallel processing logic machine."

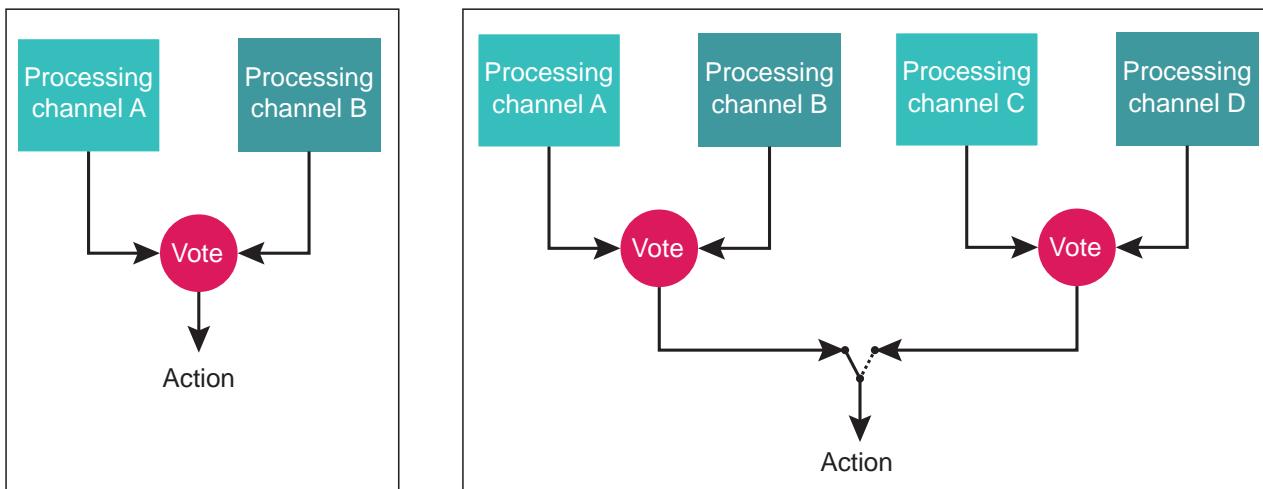
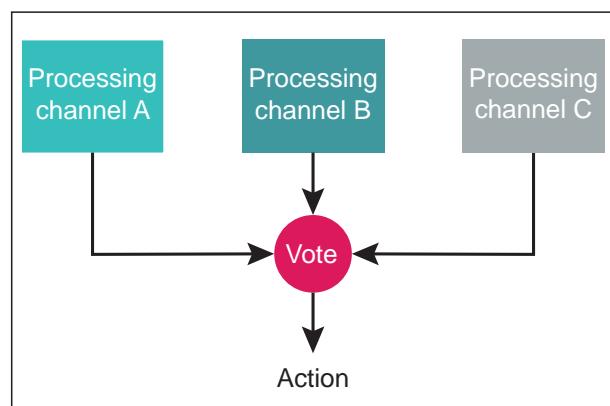


Figure 6 – Achieving safety and availability in interlocking systems typically involves the adoption of one of the architectures shown here.

Above left, 'two out of two' requires both processing channels to determine a course of action which will only be carried out if both agree. If either fails then the system shuts down to a safe state.

Above right, duplicating two sets of 'two out of two' and switching between the two pairs increases system availability.

Right, 'two out of three' uses three processing channels. At least two channels must agree on an action before it is taken, but failure of a single channel will not lead to a shut down.



necessary for an interlocking. Failure modes of microprocessor-based systems are much more complex and unpredictable than in relay logic, and their causes can be difficult to trace. These causes include electrical interference, unstable supply voltages, poor programming (leading to memory stack overflows, race conditions, deadlock, etc), derived requirement errors and manufacturing defects.

The architecture of the software-based interlocking must be designed so that the overall level of safety is at least as safe as the best relay interlocking, despite the relatively low integrity of the component parts and the complexity of their failure modes. The basic approach is to have two separate processing channels in the interlocking, each one executing the route requests from the control panel (or desk/VDU) in accordance with the signalling principles and application rules for the particular track and signalling layout. This is known as a 'two out of two' (2oo2) configuration. In the event of a difference in the outputs from the channels (indicating an error has occurred), the system shuts itself down and a safe state is enforced. In most systems lineside signals return to danger (stop) and points cannot be moved.

In practice, achieving a safe shutdown is not quite as simple as it might at first appear. Firstly, the part of the system that compares the outputs of the two channels and shuts the interlocking down in the event of a difference must be highly

dependable. A simple electronic comparator that is monitoring the two outputs is not sufficient. Secondly, there is the problem of common mode failure. Since both channels are executing the same task, there could be processor problems or programming errors which would affect both channels in the same way. In such circumstances there would be no disagreement between the outputs presented to the comparator, and the system would consequently not shut down.

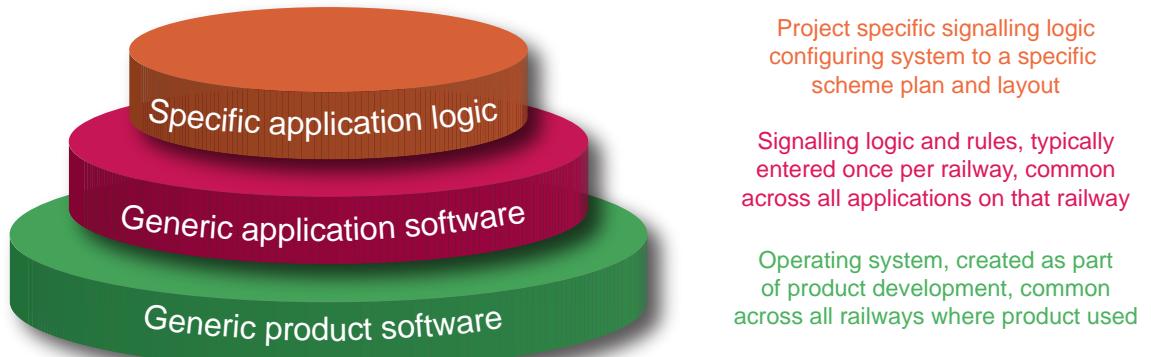
There are various solutions to these problems, of course, and different system manufacturers adopt different approaches. These may involve:

- using different hardware and/or software for the two channels to reduce the likelihood of common mode failure (often called 'diversity').
- more complex cross-checking of internal states, inputs and outputs between the two channels in order to detect faults.
- more than one mechanism by which a shutdown can be enforced (and employing special hardware for the purpose).

These mechanisms are not without their difficulties. For instance, a lack of synchronisation (differences in timing) of processing in the two channels can cause the two channel outputs to be different for short periods of time, even though each channel is behaving correctly. These short-term differences may be interpreted by any cross-checking as an error and cause a shutdown, creating a serious threat to reliability.

"Failure modes of microprocessor-based systems are much more complex and unpredictable than in relay logic."

Figure 7 – Most computer-based interlockings split their software and configuration into a number of layers, enabling the same basic hardware to be used in many different applications.



Reliability of CBIs

The reliability of computer-based interlockings is almost as important as their intrinsic safety. A fault in a relay interlocking (such as a failure of a relay) may cause a small number of routes to be inoperable, but it is very unlikely to render the whole interlocking unusable and thus stop all trains. But a computer-based interlocking that detects a processing fault may shut itself down completely, stopping all trains in the area of control. Most modern CBIs therefore have built-in redundancy to improve reliability.

One approach is to have three processing channels in the interlocking, instead of two. If one channel disagrees with the other two, a majority voting system shuts it down and the interlocking continues operating with the two remaining channels. This configuration is known as 'Two out of Three' ('2oo3') and was popular in early CBIs when computers were expensive, because it used less hardware than the alternative arrangement described below. There is a marginal safety disbenefit in this arrangement, because very rarely it could be that one channel is correct and the other two are in error. In addition, if the same software is used in all three channels then common mode failure remains a risk and producing three diverse sets of code and/or hardware to avoid such failures would be very expensive and create an even greater risk of timing issues. However, the other safeguards built into the systems makes these issues extremely unlikely in practice.

Alternatively, some interlockings have a complete duplication of the two channels (i.e., two sets, each comprising two channels) – a configuration known as 'Duplicated Two out of Two' (2X2oo2). If one set identifies a disagreement between its channels, it is shut down and the other set (which is in effect a 'hot standby') continues to operate the railway. This tends to be a more popular arrangement in modern interlockings, as it is easier to implement than 2oo3 and the cost of the additional hardware is not such a big issue as it used to be. The set of hardware and software that is acting as the hot standby must have all the same inputs and be kept in complete synchronism with the controlling set otherwise the changeover will not be seamless and some form of initialisation process will be required.

"The safety of the railway is critically dependent not only upon the core interlocking product but also upon the correctness of the specific application data/logic"

Configuring CBIs

All interlockings must be configured for the particular track and signalling layout required – a task generally performed by a signal design engineer. In the case of computer-based interlockings, he or she has to produce configuration logic (program code and/or data) – a process commonly known as 'data preparation'.

The concepts of 'free-wired' and 'geographical' relay interlockings have their equivalents in computer-based interlockings, each with their advantages and disadvantages. Both make use of the duplication and redundancy techniques described above to achieve required levels of safety and reliability.

The free-wired equivalent typically uses general purpose safety PLCs which are configured for each specific railway layout. A notation known as 'Ladder Logic' is frequently used to configure the system, although it is also possible to use fundamental Boolean logic or more sophisticated PLC programming languages. Ladder logic resembles a conventional circuit diagram that has switches, relay coils and contacts, and other electrical elements such as counters, latches and timers. It is therefore intuitively easy to produce by someone familiar with relay circuits.

The equivalent of a geographical relay interlocking uses a more conventional form of microprocessor-based computer. The core system (the 'generic product') is customised by incorporating standard software modules which define how basic track and signalling elements operate in accordance with the signalling principles for the railway on which it is to be used. This hardware and software package is the signalling manufacturer's interlocking country/client-specific product (known as the 'generic application'). The signal design engineer then produces application data which defines how the country/client-specific software modules are configured to represent the particular track and signalling layout (known as the 'specific application'). The data format is usually proprietary to each manufacturer's system.

In both types of interlocking, the safety of the railway is critically dependent not only upon the core interlocking product but also upon the correctness of the specific application data/logic, which is why so much effort goes into checking and testing it.

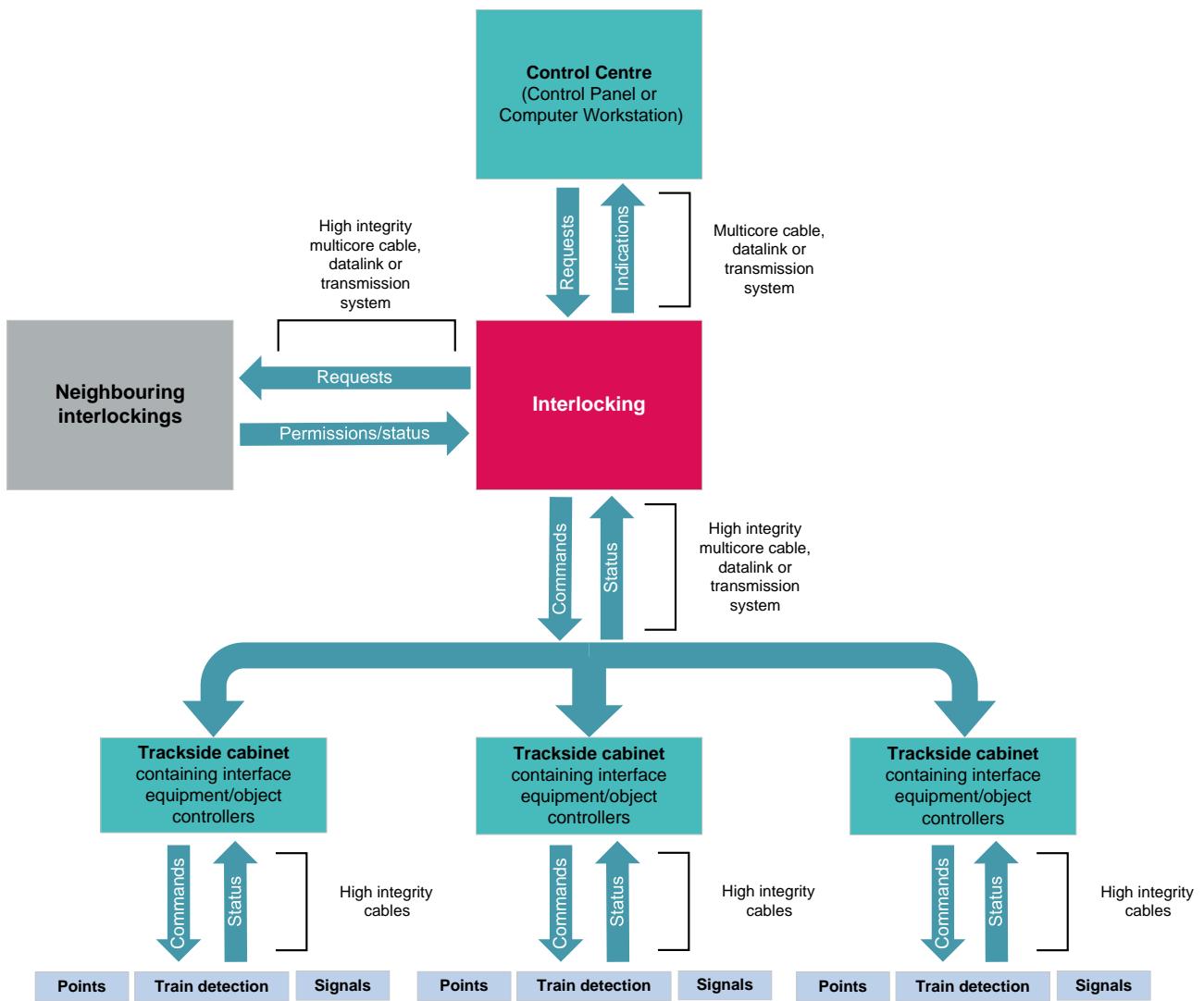


Figure 8 – The interlocking forms the ‘safety layer’ of a railway signalling system, receiving information from and relaying status to a control centre, communicating with adjacent interlockings, and controlling trackside objects.

Nowadays a large proportion of the data and logic for computer-based interlockings can be generated automatically from the signalling scheme specifications. Signal design engineers can therefore concentrate their skills on the special or unusual interlocking elements of a signalling scheme which cannot be designed automatically. Simulation and automated testing can also reduce the amount of manual verification and validation required. Because interlockings primarily implement a set of logical rules they are particularly well suited to testing and validation using formal methods and automation. Most of the major interlocking suppliers have now adopted these methods in some form and as a result the number of errors found by more conventional testing, particularly in the field, has reduced very significantly.

The interfaces between an interlocking and other sub-systems

In a modern signalling system, an interlocking interfaces with a number of other sub-systems. The three most important interfaces are with the trackside equipment, with the control panel (or computer workstation), and with other neighbouring interlockings. There are of course other interfaces, but we are not attempting here to describe the architecture of a complete signalling system, so they are not explored in this article.

Interlocking to control panel/desk interface

The signaller controls the movement of trains either by use of a control panel or by using a control desk and workstation. A panel is equipped with a representation of the track layout, on which are buttons and switches for setting routes etc, and indications to show him or her information including what routes are set, the positions of trains and the aspects of signals. In the case of a control desk and workstation, the signaller has the same information presented to him/her on screen and sets routes etc by use of a keyboard and mouse.

The interface with the interlocking is therefore two-way, with route requests being sent from the control panel/desk to the interlocking and information from the interlocking being sent to the control panel/workstation. There is usually some sort of interface sub-system (either relays or software-based) between the two.

A typical control panel/desk will communicate with several interlockings, as the geographical area covered by the control panel/desk is often larger than that covered by a single interlocking. Where all the interlockings are housed in the same building as the control panel/desk, the communication with the interlockings is achieved either by multi-core cables or by a data



Figure 9 – The use of computer-based interlocking, and increasingly the use of network communications, allows modern control systems to be implemented, as in this example from Hong Kong MTR.
The interlocking continues to ensure that the signalling rules are enforced to assure safety, but the non-vital control system above allows the railway to be operated optimally.
Photo Francis How.

"Just as with relay interlockings, the integrity of the communication link is vital for safe operation"

communications bus. If some of the interlockings are in buildings some distance from the control centre, then some form of remote control and indication system (using, for instance, time-division multiplexing) is often used to connect interlockings to the control centre to reduce the amount of cabling required.

It is important to note that all these communication links are not safety-critical (i.e. they do not have to be SIL 4). The interlocking ensures the safety of train movements even if there is a fault or failure in the communications links or interface sub-systems. That said, the interfaces and communications links must be as reliable as possible, both for normal working and in degraded mode situations. In both cases, the control panel or workstation is presenting the signaller with information about the railway, upon which he or she makes decisions regarding the movement of trains.

Many modern signalling systems include Automatic Route Setting (ARS), a sub-system which automatically sets routes ahead of each train based on the timetable and, where conflicts arise (eg because of late running), on a set of rules for prioritisation of train movements. This relieves the signallers of much of their routine work. So far as the interlocking is concerned, however, it receives and acts on ARS route-setting requests in the same way as if the requests had come from the signaller and the control panel/workstation.

Interlocking to trackside equipment interface

In the case of a relay interlocking, in most systems the items of trackside equipment (signals, points, track circuits etc) are connected to the interlocking by multi-core cables, with a dedicated pair of cores for each circuit. The cables are generally specified and constructed to meet railway requirements, both in terms of resilience to the trackside environment, and in terms of the integrity and separation of each core. This is necessary because they are carrying safety critical

circuits for operating the signals, moving points, indicating the occupancy or otherwise of a track circuit to the interlocking etc. The interlocking depends upon the integrity of these circuits in order to function safely.

In the case of computer-based interlockings, the items of trackside equipment are normally (but not always) connected to nearby object controllers which provide the power for the equipment, pass the interlocking commands to the equipment and receive equipment status information for sending back to the interlocking. The object controllers are connected to the trackside equipment by high integrity cables as described above, and to the interlocking via a communications datalink. Just as with relay interlockings, the integrity of the communication link is vital for the safe operation of the interlocking. The object controllers therefore generally have a 2oo2 configuration to ensure safety, and the datalink uses highly secure coding techniques to prevent (or detect) corruption of the transmitted data. In some manufacturers' systems the communications protocol is proprietary to their product, but increasing use is now being made of IP addressing techniques. Manufacturers still use their own safety and applications protocols, although the European EULYNX project is promoting open standards for interfaces, to reduce signalling life cycle costs.

Interlocking to interlocking interface

Interlockings must be able to interface with neighbouring interlockings, because almost inevitably at the geographical boundaries there will be routes that have their entrance signal in the area controlled by one interlocking and their exit signal in the area controlled by another. The route setting process is initiated by the interlocking responsible for the entrance signal, but requires action by, and information from, the other interlocking in order for the complete route to be declared 'set', before the entrance signal is allowed to clear.

It is common practice, so far as possible, to arrange the geographical boundaries of interlockings to occur on sections of plain line, where the interlocking arrangements for each cross-boundary route are very straightforward. However, this is not always possible, particularly in places such as complex station areas where more than one interlocking is required. In the case of relay interlockings, the interface generally takes the form of high integrity multi-core cables to link circuits in the two equipment buildings – the same sort of cables as are used to connect the relay interlockings with their items of trackside equipment.

In the case of computer-based interlockings, a high integrity datalink is used for the interface (bearing in mind that the two interlockings may well be in physically adjacent cabinets in an equipment room rather than in separate buildings). However, the interfacing arrangements tend to be more complex than in the case of relay interlockings. The two sets of software have to act together and perform 'handshakes' with each other in order to set, lock and release routes.

The future of interlocking

At the beginning of this article we said that 'interlocking' is the mutual dependency between signalling functions (moving points, clearing signals etc). We have seen how these interlocked functions are made real by use of some sort of 'logic machine', whether mechanical, electrical or computerised, which signal engineers call an interlocking. In a modern signalling system we might expect to see a relay interlocking or a suite of computer interlockings housed in a building or in a cabinet trackside, connected to trackside equipment, the control panel/desk and screen, and to other interlockings by cables and data transmission systems. But the concept of a physically discrete interlocking is starting to change.

Various examples come to mind. Firstly, in the case of ERTMS Levels 2 and 3, there is a second vital element in the system, namely the Radio Block Centre (RBC), which links trains with the interlocking, sending information about movement authorities, permissible speeds etc and receiving information about the train speed and location. Like the interlocking, the RBC must be of the highest safety integrity. System suppliers are starting to combine these two functions on the same computer platform, and in the world of

"The concept of a physically discrete interlocking is starting to change"

About the author ...

Francis has been a long-time member of the IRSE. First with British Rail/Railtrack, Atkins, as the technical director of the Railway Industry Association and chief executive of the IRSE. He was an IRSE exam Thorrowgood scholar and served on Council for many years and was president of the Institution 2012-2013. He is widely respected for his professionalism and

metros, a number of systems already combine the functions of interlocking and track-train messaging within the same system.

A second possibility, and one that is already being explored, is to use cloud technology for interlocking. This has various advantages, including cost, flexibility and, potentially, resilience. It also comes with challenges, not the least of which is maintaining cyber-security (and therefore safety as well as reliability).

Thirdly, we may see a move towards distributed interlocking functionality. In recent years the trend has been to place computer-based interlockings together in a single location (often co-located with the control centre). In time this may change, with some interlocking functions shared between the train-borne and trackside systems. Indeed, it could be argued that this is already happening to a limited extent with CBTC and ERTMS.

Even more radically, the train may play a key role in initiating the setting of the route ahead and in determining its own safe movement authority, making use of train-to-train communications to do so. A more train-centric architecture will be adopted, with the trains being 'smart' rather than simply responding to movement authorities issued by trackside infrastructure systems. Again, suppliers are already starting to make some of this a reality. This may well lead in time to some of the traditional principles of signalling being challenged, such as making the distance between following trains dependent upon their relative speeds rather than always assuming the train ahead is stationary.

Closing remarks

This article has provided an introduction to the technologies used for railway interlocking. Next month we will look at the functionality of an interlocking.

If you want to know more, some of the IRSE textbooks cover the subject in greater detail. For many signal engineers the specification and design of interlockings is at the heart of their careers. It requires knowledge, experience and expertise – and it is vital to the safety of the railway. But if you are new to the industry, don't let that deter you. Instead, take every opportunity to learn from those who have the experience and knowledge.

technical knowledge and played a vital role in drawing younger members into the running of the Institution and has encouraged and helped them develop their capabilities in both their professional and IRSE roles. He has given quiet encouragement and encouraged self-confidence in many of the rising engineers in the control and communications industry.

Back to basics: Interlocking

Part 2



Francis How



Last month in IRSE News there was the third in a series of articles on 'back to basics' themes, looking at the essentials of 'interlocking', focused on the technology used. This month we are going to look at the functions that an interlocking performs, and how these ensure the safe movement of trains.

This article is no more than an introduction to the subject, intended for IRSE members new to the industry rather than those who are experienced in specifying, designing and testing signalling systems.

The functions of a modern interlocking

We saw in last month's article that interlockings have used a variety of mechanical, electrical, electronic and software-based technologies over the years. But regardless of what technology is used, a route-setting interlocking must perform essentially the same functions in order to ensure the safe movement of trains. These functions, which are defined in the railways' signalling principles and application rules, must be compatible with the operational rules/regulations for the movement of trains, under both normal and failure conditions.

The signalling principles and application rules for the interlocking functions vary somewhat from country to country but are similar in their basic requirements. The descriptions of the principal functions in this article are based on the current signalling principles used for colour light lineside signalling of passenger railways on the mainline railway in Great Britain and in some other countries (they are relaxed somewhat for shunting, permissive and freight movements, but these are not covered in this article). Where some railways adopt significantly different practices, these are noted but are not described in any detail.

"Signalling principles vary from country to country but are similar in their basic requirements"

Checking route availability

When a request to set a route is sent from the control panel/desk/VDU (or from the Automatic Route Setting system, where one is provided), the interlocking first checks that the route can be set in its entirety, and that the request does not conflict with any other routes that have already been set or are in the process of being set. Without this check, the interlocking might start moving points to set the route but fail to complete the process because one or more points are locked in the wrong position by routes that are set for other train movements. The check is also vitally important for 'directly opposing' routes, for trains travelling in the opposite direction to the route whose availability is being checked, and for which the point settings are identical.

If the whole of the route is not available for setting at the time the request is received by the interlocking, it is rejected or ignored, rather than being stored until the route can be set. This is a feature known as 'anti-preselection'. Not all railways include it, but it is often regarded as good practice to prevent a route request from continuously trying to make the interlocking set a route that cannot (yet) be set.

Route setting and locking

When the route availability check is successfully completed, the interlocking starts moving the points in the route to their required positions. It may also be necessary to move other points that are not in the route itself, to protect the route from other trains in the event that they pass a signal that is displaying a stop aspect.

When the points are correctly positioned, route locking is applied to all the track detection sections that form the route, thereby reserving each section for the route being set, preventing the points from being moved and ensuring that no conflicting routes can be set.



Interlocking has been an underpinning concept for railway safety since the 1850s, although the technology upon it depends has changed phenomenally during that period. Areas of points and crossings are the most complex part of any interlocking design.
Photo Paul Darlington.

"Some railways don't use overlaps at all"

Overlaps

Many countries and railway administrations also lock and protect a short section of track beyond the exit signal as part of the route-setting process. This is called the 'overlap', and it is typically between 50m and 200m in length. The provision of overlaps is not a universal practice, however. Some railways do not use overlaps at all and, at the other extreme, some have overlaps which comprise, in effect, all the track from the exit signal to the next signal beyond that. Even on railways that use overlaps, they are not necessarily required for all types of routes. Where railways do use overlaps the rules for setting and locking them vary from one railway administration to another, so it is worth emphasising that this description applies to main line railways in Great Britain and is not necessarily true of other railways.

The purpose of an overlap is that if the exit signal is at danger (stop) and the train fails to come to a stand at the signal because of inadequate deceleration, it is likely to stop within the overlap distance and so avoid risk of collision with other legitimate train movements. Of course, it is not guaranteed that the train will stop within the overlap, and an overlap is of no help at all if the train brakes have not been applied or the adhesion conditions are very poor. It should also be noted that with cab signalling, overlaps have an additional purpose, to do with the accuracy with which the train location is known.

Trailing points in an overlap are set to the correct position and locked when a route is set. Where there are facing points ahead of the exit signal there may be more than one overlap that can be used. These facing points must be set to the position required for the selected overlap, and they may also be locked. The reason for not locking the facing points in all circumstances is that in complex areas such as stations, the interlocking may permit the overlap to be changed after the entrance signal has been cleared (a

feature known as 'swinging overlaps') by moving the position of the facing points to create a new overlap. This provides signallers with greater operational flexibility. The facility to swing an overlap is inhibited by the interlocking as the train approaches the exit signal so that a safe (locked) overlap exists in case the train passes the signal at danger. Designing the circuits or data for swinging overlaps can become very complex indeed, however, and their provision should be limited to that which is considered necessary for operational purposes.

The setting of an overlap happens only if the exit signal is at danger (stop), of course. If a further onward route is set for the train, from the exit signal to the next signal beyond, then that route is itself set and locked in the same manner as described above.

Clearing the entrance signal

When the route has been set (as described above), all relevant train detection sections must be proved 'clear' (i.e. no train or vehicles present) before the entrance signal is permitted by the interlocking to show a proceed aspect. This includes:

- All the track that forms the route between the entrance and exit signals.
- All the track that forms the overlap ahead of the exit signal (if an overlap is provided).
- Any other sections of track on which a train or individual rail vehicle could stand and be 'foul' of the route (i.e. with which the authorised train could collide).
- Any 'flank' sections of track, which are included to provide early detection of another train passing its own exit signal at danger (SPAD) and thereby intruding onto the route of the authorised train.

In the case of lineside signalling, other checks may also be required before the signal is permitted



Interlocking design is based on the avoidance of conflicting routes, allowing for a wide range of combinations of possible events.
Photo Paul Darlington.

"There may be other requirements that have to be satisfied before the interlocking will permit the entrance signal to clear"

to clear, including the following. Not all railways apply these checks, and the specific rules for applying them vary from one railway to another:

- The exit signal must be alight (i.e. displaying a visible aspect), to avoid the risk of the driver failing to see it.
- If the entrance signal is required to display a route indication for the route set, that indication must be proved alight before the entrance signal is permitted to display a proceed aspect.
- If the train is being routed over a diverging junction ahead of the entrance signal, that signal may be held at a restrictive aspect (stop or caution) by the interlocking until the train has slowed down sufficiently for the diverging points and the route beyond.

There may be other requirements that have to be satisfied before the interlocking will permit the entrance signal to clear, for instance for level crossings or train protection systems (the latter to stop, or mitigate the risk of, the train passing its exit signal at danger).

The actual aspect displayed by the entrance signal when it clears, including any speed or route indications, depends not only upon the route immediately ahead, but also on whether a further route has been set for the train beyond the exit signal. The sequence of aspects seen by a driver at successive signals as he or she approaches a signal at danger (stop) varies from one railway to another and is not dealt with in this article.

On most railways that use colour light signalling, the entrance signal will revert to danger (stop) if the conditions that permitted it to clear are no longer fulfilled. So, for instance, if a track circuit in the route fails, or there is a loss of detection on a set of points, the entrance signal will automatically revert to danger. This is a safe arrangement but can be worrying for a driver who unexpectedly encounters a signal at danger (stop). Some railways do not include reversion – and of course, with mechanical signals there was no possibility of doing so!

Route holding and release

When the train passes the entrance signal (showing a proceed aspect), the interlocking returns the entrance signal to danger (stop) but maintains the locking ahead of the train to prevent points from being moved and conflicting routes from being set. In order to maximise capacity and flexibility it is, however, desirable for the locking to be removed as soon as it is safe to do so after the passage of the train, in order that other routes may be set which make use of the same track (or some of it).

Before the locking can be released after the passage of the train, the route request must first be cancelled. Traditionally this would be done by the signaller using the control panel/desk. However, to ease the signaller's workload and to enable the locking to be released as soon as possible, many modern interlockings include a function called 'train operated route release' (TORR), which cancels the route request from the control panel/desk when the train has passed the entrance signal, without any action by the signaller. The interlocking normally does this by checking that the first two or three train detection sections beyond the signal show 'occupied' and then 'clear' in the correct sequence with the passage of the train (this minimises the risk of a train detection failure leading to the premature release of the route). Although the route request is cancelled by this process, the route locking is maintained to ensure the safe passage of the train.

The interlocking may also have a function which permits individual portions of the route to be unlocked as soon as possible after the train has passed, rather than having to wait until the train has passed through the whole route. This is known as 'sectional route release'. As the train passes clear of each train detection section in the route, the locking on that section is released, provided that:

- The entrance signal has returned to 'stop' (danger).

- The route request has been cancelled (either by the signaller or by TORR).
- All the route locking from the entrance signal up to the start of the section has already been released.

Route locking ahead of the train continues to be maintained.

Approach Locking

There is a further category of route release, applicable only where the route has to be cancelled **before** the train passes the entrance signal. Clearly this is an unusual set of circumstances – for instance if the signaller needs to change the order of two trains at a junction after the route has been set for one of them, or if there is an emergency and it is necessary to try to stop the train. In such circumstances the signaller cancels the route on the control panel/desk, which has the effect of setting the entrance signal to stop (danger). This action does not necessarily immediately release the route locking ahead of the entrance signal, however, in case the train cannot brake sufficiently to stop at the signal and consequently enters the route beyond it.

The interlocking function that determines when the route locking ahead of the entrance signal is released in the circumstances described above is known as 'approach locking'. In its more comprehensive form, it maintains the locking of the route ahead of the entrance signal (which is displaying stop/danger) until one of the following conditions is satisfied:

- The train has come to a stand at or before the signal, or
- The train has sufficient braking distance to come to a stand at or before the signal.

With lineside signalling the first of these conditions is usually achieved by using a timer in the interlocking, rather than directly confirming the train is at a standstill. When it has finished timing, the train is assumed to either have come to a stand at the signal or to have passed it, being unable to stop in time (in the latter case the route locking holds the route safe for the train). The second condition consists of a check by the interlocking that the train has not yet occupied any of the train detection sections between

"There are many other functions that may feature in an interlocking"

Interlocking design has to consider many factors including the speed of the trains using the railway and interfaces to other equipment such as level crossings.

Photo Shutterstock/
VanderWolf images.

the sighting point of the first signal displaying a caution aspect and the entrance signal displaying stop. With cab signalling, the speed and location of the train are usually known and can be used to check the two conditions, which is a more accurate method of checking whether the conditions are fulfilled.

Releasing the overlap

If the exit signal is at danger (stop), and the train has safely stopped at it, then it is necessary to release the locking of any points in the overlap beyond the signal. The reason for this is that the points may either not be set correctly for the train's onward route, or because another train needs to use some of the track and points in the overlap.

With lineside signalling, the interlocking times the train's occupancy of the last train detection section on the approach to the signal. When the timer finishes, and provided that the train detection section immediately beyond the signal has not been occupied, the interlocking will release the points in the overlap. With cab signalling, the speed and location of the train are usually known and can be used directly to check the train is at a standstill at the signal.

When the interlocking determines that the train is stationary at the exit signal, the locking of the overlap is released and the track and points in it can be used for setting the onward route, or for routing other trains.

Other interlocking functions

This article describes only the basic interlocking functionality of a signalling system. There are, of course, many other functions that may feature, including:

- The role of the interlocking in displaying speed information via lineside signals. This is of particular importance on 'speed signalled' railways, as distinct from 'route signalled' railways.
- Interlocking the signalling system with protection systems for personnel working on the track.



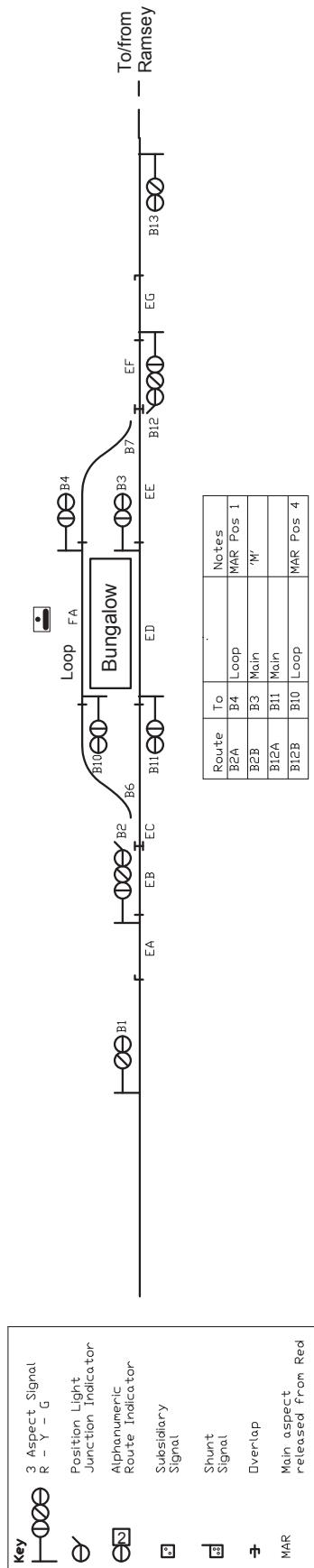


Figure 1 – Simple Scheme Plan, showing the signals, train detection sections, points, overlap etc. This example was used in a recent IRSE examination.

Route Information	Points Free to Go or Release	Route Normal	Opposing Route Locked Locked after route set until	Points Called, Locked and Detected	Tracks	Aspect Sequence	Approach Locked when	Approach Released	Special Controls
Entrance signal / Route	Class	Normal Reverse	Tracks Clear	Normal or Reverse	Proved Clear Proved occupied for	Aspect	Indication	Signal Cleared/ Route set	First Condition Second Condition or Time (secs)
B12A	B11	Main	B7, B6	-	B3 EE B6 #10	EE #11, ED, EC	Y	-	EF cll. EE occ
Assumes									
<ul style="list-style-type: none"> Modern Network Rail practice, including: <ul style="list-style-type: none"> TPV/S train protection designed to stop train within overlap length Route locking incorporates track bob protection Sufficient standage in the loops that the longest trains may pass; therefore special controls are not necessary. No special controls relating to the method of single line working are necessary (method is unspecified on the plan but would appear not to be Track Circuit Block, Token Block nor One Train Working- No Staff). 									
#10 Point set and locked, but not detected #11 Stick track - returns signal to red (danger) when occupied when signal off and berth track also occupied #12 Only applies when Temporary Approach Control Facility selected by technician #13 Includes proving that associated TPV/S loops energised #14 AWS electro live when signal at green and alight, AWS suppressor energised after route set from B3 or B4 provided that no Up direction movement from Ramsey is possible. #15 TPV/S energised whilst signal controlled to Red #16 Route cancelled by triggering of Pre-Defined Overrun Protection Group by a SPAD detected at B2 #17 Route cancelled by activation of Signal Group Replacement Control									

Figure 2 – Example of a signal/route control table for the route from signal B12 to B11 (as shown in Figure 1). Our thanks to Peter Woodbridge for creating this example.

However functional and complex the interlocking is, it's important to ensure that the information relayed to the driver is clear and easily understood, as an aspect indicating safe speed (left) or route based signalling.

*Photos Shutterstock/
Miles Schofield and
Paul Darlington.*



- The provision of train protection systems (which will be covered in another 'Back to Basics' article, soon to be published in IRSE News).
- Interlocking with level crossings or other moveable infrastructure (e.g. swing bridges).

Specifying interlocking requirements

The process for specifying the interlocking requirements for a particular track layout is strictly governed by railway administrations (or in some cases their regulatory bodies), because of the safety implications of an error.

Ideally, the starting point for the signal engineer is to be provided with details of the proposed track layout and the operational requirements. The first of these includes knowledge of the location of points and associated critical dimensions. The operational requirements include the frequency of trains, their maximum speeds and braking capabilities, the required headways and the specific train movements (main line, shunting, permissive etc).

In practice, this ideal starting point is not always the reality! Signal engineers may, for instance be told to replicate the existing signalling arrangements in modern form (which is itself an ambiguous statement), without being given any explicit statement of operational requirements.

The signal engineer also needs to know the signalling principles applicable to the railway for which the design is being prepared. These are the high-level generic rules for ensuring the safe movement of trains, and they cover all the requirements relating to route setting, locking and release – and more.

The knowledge of the track layout, the operational requirements and the signalling principles enables the signal engineer to produce two key sets of documents. The first is the Scheme Plan, and the second are the Control Tables.

Scheme Plan

The Scheme Plan depicts the layout of the track, showing points, signals, train detection sections, level crossings, stations, permissible speeds etc. Each track is usually shown as a single line, not a pair of lines. Each signalling object is allocated a unique identity (numbers and/or letters). The plan also shows the routes that each train can take from each signal. Each route is given a unique identity. Other relevant information may also be shown on, or be associated with, the Scheme Plan. A simple Scheme Plan is shown in Figure 1, for a single line with a passing loop at a station. The symbols shown are used by many railways around the world, but are not universal.

Control Tables

The Control Tables state all the conditions that must be fulfilled for each route before a train can be given a movement authority to use it. The route information is derived from the Scheme Plan – the position of points, the train detection sections that must be clear, other routes that must not be set etc. Approach locking release conditions are also specified, as are requirements for functions such as the delayed clearance of signals at diverging junctions, and the aspect sequence for successive signals. A simple example of a signal/route control table is shown in Figure 2, for a route from signal B12 to B11 on the Plan in Figure 1. The requirements are based on current British (Network Rail) signalling practice and include some features (such as train protection proving and technician controls) that we have not dealt with in this article.

Control Tables are also prepared for each set of points, stating the conditions that must be fulfilled before a set of points is free to move. This includes the routes that require the points to be normal or reverse, dead-locking train detection sections and point-to-point locking.

It is in the preparation of the Control Tables that the signal engineer's knowledge of the signalling principles and application rules is of vital importance. From the Control Tables, the circuit diagrams for relay interlockings (based on standard circuits) can be produced, or the data in the case of a computer-based interlocking.

The production of the Control Tables is, therefore, a critical step in the design and configuration of the interlocking. Nevertheless, the gradual automation of the design process means that it is possible to go directly from Scheme Plan to detailed design, with the Control Tables being a by-product of the process for later use (e.g. for recording tests performed on the interlocking), rather than being a key stage in the design process.

Closing remarks

This article and last month's have provided an introduction to railway interlocking – and if you want to know more, some of the IRSE textbooks cover the subject in greater detail. For many signal engineers the specification and design of interlockings is at the heart of their careers. It requires knowledge, experience and expertise – and it is vital to the safety of the railway. But if you are new to the industry, don't let that deter you. Instead, take every opportunity to learn from those who have the experience and knowledge.