Logical Verification and Implementation of Protocols
G.V. Bochmann
Proceedings of the Fourth Data Communications Symposium, Quebec, Canada, 1975
Pages 8-5 — 8-20
© 1975 by the IEEE. Reprinted by permission.

Abstract
The implementation and logical verification of communication protocols are considered in view of obtaining reliable communication systems. It is proposed that methods for specifying communication protocols should be useful for obtaining a comprehensive description, and for simplifying the logical verification and implementation of the specified protocol. These issues are discussed for a particular protocol, which is taken as an example throughout the paper. The example chosen is a simple HRC protocol for two-way simultaneous data communication over a point-to-point data link. After an informal description, its operation is defined precisely. From the definition of the protocol, certain properties are proved which show the correct operation of the protocol. In particular, it is shown that line transmission errors are correctly recovered, and that the sequence counts can be represented modulo a given constant. A condition for completed data transmission is also given. The method for proving these properties is general, and can also be applied in the case of the more complex protocols that are used in actual applications. The second part of the paper deals with the problems of parallel processes and the implementation of protocols. Well structured high-level languages which include facilities for describing parallel processes and monitors are proposed as tools for the implementation of protocols. As an example, it is shown how such a language can be used to program the protocol introduced earlier. It is pointed out that this program can be easily obtained from the original protocol definition by doing only a few changes and refinements. Therefore it is expected that such an implementation is relatively reliable.

Introduction
The design and implementation of appropriate communication protocols is an important part of the development of any new communication or distributed computer system. Normally a protocol is designed to operate efficiently in normal situations, and also to deal with occasional erroneous behavior or malfunction of subsystems. The different situations a communication protocol must cope with, are well described in references 1. A real implementation must, however, take into account some other important properties. A protocol should be designed with certain features in mind: (a) Normal communication between subsystems. (b) Occasional erroneous behavior of one or several of the subsystems. This includes the occasional malfunction of the communication line. The recovery procedure will normally consist of trying the same action again. (c) Long range erroneous behavior or failure of one or several subsystems. The recovery procedure will normally consist of reconfiguration of the system.

The design of a communication protocol is a difficult task because many different situations must be considered. Besides the normal operation of the protocol, situations due to erroneous behavior of one or several subsystems must be described. The synchronization between the subsystems in these situations is not very simple to understand. Therefore a logical verification of the protocol would be very useful. By logical verification we mean proving certain properties of the protocol which assure its correct operation in all situations. After its design, a protocol must be implemented in software and/or hardware. An implementation must be easy to understand, correct and efficient. The choice of a good programming language for the implementation is of large importance. The method used for specifying protocols has a strong influence on how easy or difficult it is to design and implement a protocol for a given communication problem. We believe that the design, logical verification, and implementation of a protocol are interrelated. Therefore a method for specifying protocols should be useful for all three of these activities. More precisely, a method for specifying protocols should have the properties that:

1) a protocol can be specified in a comprehensive form; in particular, the complete definition of a protocol can be partitioned into different levels of abstraction;
2) the specification of a protocol allows proving certain properties of the protocol and its operation, proving in particular that the error recovery is effective, and that all possible situations of erroneous behavior have actually been considered;
3) given the specification of a protocol, its implementation is simple, and part of the implementation may be obtained automatically.

It seems that at present, there is no method for specifying communication protocols that satisfies all these requirements. However, several tools for specifying, proving properties of, and implementing systems of parallel processes have been discussed in literature. The purpose of this paper is to show by an example that certain tools that have been developed in different fields of computer science can be useful for the design and implementation of communication protocols. We discuss in particular the possibility of logical verification by proving certain properties of protocols, and implementing protocols using a high-level language for parallel processes. Logical verification of protocols is useful for obtaining reliable communication systems, because it helps detect weak points or errors in the protocol design which are difficult to find by simulation or testing. The use of a well-structured high-level language for the implementation of protocols has many advantages. It simplifies the programming effort, facilitates the detection of programming errors, makes the system more transportable, and clarifies the program documentation.

In this paper, we use a simple protocol for two-way simultaneous data communication as an example. This protocol is informally introduced in section 2. In section 3, we give a formal specification of the same protocol. This specification is then used to derive certain properties of the protocol and its operation. In particular, we show that the protocol correctly transmits message sequences, and that the internal sequence counts can be represented modulo a certain constant, thereby allowing a reasonable implementation. We use the technique of assertions for verifying these properties. We note that this technique cannot be used for proving the absence of undesired loops, and the effective termination of a transmission sequence. These problems, however, have been considered in references 4.

In section 4, we discuss how the concepts of parallel processes and monitors 5-7 can be used for the implementation of a protocol in a high-level programming language. As an example, we give the protocol
COMMUNICATION PROTOCOL MODELING

3. A proof of correct operation

The section above gives an informal description of a simple HDLC protocol. In the following we show how certain useful properties can be proved about its operation. In order to do this, we first give a more formal definition of the protocol. This definition also serves as starting point for obtaining an implementation of the protocol, as discussed in section 4. From the definition of the protocol, we prove certain assertions about the states through which the protocol can pass during its operation. These assertions are then used to prove that the protocol correctly recovers the transmission errors of the data link, and that the sequence numbers can be represented modulo the constant 2M.

3.1 Definition of the protocol

The information to be transmitted from a source to a sink consists of a sequence of messages. Each message is transmitted as the information field of a single frame (i.e. transmission block). We write the i-th frame that travels on the transmission line in the form $\langle S_1, R_1, n_1 \rangle$, where $S_1$ and $R_1$ are the send and receive sequence counters, and $n_1$ is the information field, i.e. a message; all other fields of the frame (including the redundancy check field) can be ignored for our present purposes. At a given instant, during the operation of the protocol, a certain number of frames travel on each line. The sequence of frames is written as:

$$\langle S_1, R_1, n_1 \rangle \cdots \langle S_i, R_i, n_i \rangle \cdots \langle S_m, R_m, n_m \rangle$$

where $\langle S_1, R_1, n_1 \rangle$ is the last frame sent by the transmitting station, and $\langle S_m, R_m, n_m \rangle$ is the next frame to be received by the receiving station.

Each station contains a message buffer send-buffer for the messages to be transmitted. For simplicity we suppose that its size is unlimited; later, however, we show that a cyclic buffer for 2M messages would be sufficient. Each station also contains the internal sequence counters $L$, $S$, $N$, and $R$. The meaning of the values of these counters is the following:

1. $L$ is the sequence count of the last message obtained from the source.
2. $A$ is the highest sequence for which a correct reception, at the opposite station, has been acknowledged.
3. $S$ is the sequence count of the last frame sent.
4. $N$ is the highest value that has been reached by $S$ so far.
5. $R$ is the sequence count of the last received frame that has been passed on to the sink. We note that this counter refers to the transmission in the opposite direction.

So far, we have described information transfer from a source to a sink (see figure 1, upper half). However, the protocol supports two-way simultaneous data communication as indicated in figure 1. The stations at both sides of the data link have the same structure. Each contains a sending and a receiving part. For distinguishing the two sides of the data link, the elements of the station on the opposite side (i.e. on the right) are written with an overlining bar.

![Figure 1: Two-way simultaneous data communication](image-url)
State transitions can occur due to the events of sending or receiving a frame, or obtaining a new message from the source. The state transformation of each of these events, which exclude one another in time, is described by the statements below.

We note that $M > 1$ is a system constant which determines after how many outstanding non-acknowledged frames retransmission is to start. Normally this constant will be adjusted to the expected delay for receiving an acknowledgement for a given frame, and the number of frames sent in unit time. The statements of the events are written in a free style Pascal, and comments are given in brackets {}.

(a) The event of obtaining a new message from the source
(1) obtain (message);
(2) L := L + 1;
(3) send-buffer [L] := message;

(b) The event of sending a frame
(4) if S > 1 (no new message to be transmitted) or $S > A + N$ (too many outstanding frames)
(5) then if $A = L$ then $S := A$
(retransmit last frame)
(6) else S := S + 1;
(send next frame)
(7) $N := \max(N,S)$;
(increment N if necessary)
(8) send-frame (<S,R, send-buffer [S]>); 

(c) The event of receiving a frame
(11) listen-to-next-frame (in <F Frame, frame>, message>);
(12) if redundancy-check-is-valid then begin
(13) if $F_{\text{frame}} > A$ then $A := F_{\text{frame}}$
(adjust A)
(14) if $S_{\text{frame}} = R + 1$
(correct sequence)
(15) then begin
(16) and sink-is-ready-to-receive
(17) then begin
(18) R := R + 1;
(19) pass-on-to-sink (message)
(20) end
(21) end

Property (2): At any given instant, at least A messages from the source have been correctly passed on to the sink at the opposite station.

Property (3): When $A = L$, all messages obtained from the source (and no more) have been correctly passed on to the sink at the opposite station.

We note that the assertion of correct transmission relies on the assumption that the redundancy check on received frames detects all transmission bit errors.

The properties (1) and (3) show the "correct operation" of the protocol. The following property (4) is important for an implementation of the protocol.

Property (4): The sequence counts $S$ and $R$, in the internal counters of the stations as well as in the transmitted frames, can be represented modulo 2M.

We note that similarly the counters $L$ and $A$ can use a representation modulo a constant. This constant depends on the size of the message buffer, and should at least be equal to 2M.

For proving these properties, it is useful to consider the following assertions which hold at any instant between the occurrences of the events (a), (b), or (c) defined above.

Assertions on the internal counter values (see for example figure 2):
(a) $N < S < L$ (b) $S < N$ (c) $N < L$
(d) $N = N + A$ (e) $A < R$ (f) $R < N$

For proving these properties, it is useful to consider the following assertions which hold at any instant between the occurrences of the events (a), (b), or (c) defined above.

Corresponding events can also occur at the opposite station. The order in which these events occur is not specified, however, they never occur simultaneously (mutual exclusion). We assume that there will always be a successive sending and receiving event after a finite time interval, i.e. the operation of the protocol will never stop.

3.2 Properties of the protocol

The protocol defined above transmits messages in correct sequence from source to sink simultaneously in both directions, and it recovers line transmission errors, as long as the line does not break down permanently. This can be summarized by the following properties, which hold for each of the two transmission stations at any instant between the occurrences of the events (a), (b), or (c) defined above, and which are proved below.

Property (1): At any given instant, exactly $R$ messages, from the source at the opposite station, have been passed on to the sink correctly (i.e. without any bit error, and in the correct sequence).

Property (2): At any given instant, at least $A$ messages from the source have been correctly passed on to the sink at the opposite station.

Property (3): When $A = L$, all messages obtained from the source (and no more) have been correctly passed on to the sink at the opposite station.

Property (4): The sequence counts $S$ and $R$, in the internal counters of the stations as well as in the transmitted frames, can be represented modulo 2M.

For proving these properties, it is useful to consider the following assertions which hold at any instant between the occurrences of the events (a), (b), or (c) defined above.

Assertions on the internal counter values (see for example figure 2):
(a) $N < S < L$ (b) $S < N$ (c) $N < L$
(d) $N = N + A$ (e) $A < R$ (f) $R < N$

We have so far assumed that the sequence counts can grow indefinitely. We see now that a representation modulo 2M can be used:

Property (4): The sequence counts $S$ and $R$, in the internal counters of the stations as well as in the transmitted frames, can be represented modulo 2M.

We note that similarly the counters $L$ and $A$ can use a representation modulo a constant. This constant depends on the size of the message buffer, and should at least be equal to 2M.

For proving these properties, it is useful to consider the following assertions which hold at any instant between the occurrences of the events (a), (b), or (c) defined above.

Corresponding events can also occur at the opposite station. The order in which these events occur is not specified, however, they never occur simultaneously (mutual exclusion). We assume that there will always be a successive sending and receiving event after a finite time interval, i.e. the operation of the protocol will never stop.
null
such protocols, using the same techniques.

4. Programming a Protocol

In this section we discuss aspects of protocol implementation; we suppose in particular an implementation in software. As mentioned earlier, we believe that the implementation of a protocol is closely related to its design and logical verification. Writing a program that implements a given protocol on a given computer may introduce new errors (programming errors). Therefore, it would be advantageous to use a high-level implementation language so that only few transformations are necessary for obtaining the program from the original protocol definition. Using a high-level language and structured programming are likely to increase the reliability of the protocol implementation. We discuss in the following more specifically the description of parallel processes as they occur in protocols.

The concept of parallel processes is a useful structuring tool for the design of protocols. For example, the protocol introduced in section 2 may be understood as consisting, at each station, of the following processes: (1) a message producing source process, (2) a sending process, (3) a receiving process, and (4) a message consuming sink process. Programming tools for specifying inter-process communication and synchronization have been described in the literature 5,6. The concept of monitors is particularly interesting. Monitors can be used to delimit the inter-process interaction and to specify explicitly the synchronization between the processes. They provide an efficient implementation tool 7,8.

The use of different levels of abstraction for the description of protocols is another useful design method. Often discussed in the literature on structured programming, this approach leads to the design of protocols in distinct levels 1,9. For example, the protocol introduced in section 2 could be described at a certain level of the whole communication system. The levels below concern, for instance, the redundancy check for error detection, or the modem line interface. The levels above may concern message routing in a communication network, or the implementation of mechanisms for the communication between a resource and a "user".

Independent of these different levels in the design of a communication system, the system may be described in more abstract or more detailed terms. For example, the description of the protocol in section 3.1 is relatively abstract. For an implementation of this protocol we would look for a relatively detailed specification, for instance in the form of a system written in some programming language for a given computer. In the remaining part of this section, we consider the protocol defined in section 3.1 as an example, and show how one can obtain for the same protocol successively more detailed descriptions, which lead to an implementation. We use in particular the concepts of parallel processes and monitors, and the notation of the programming language Concurrent Pascal 10.

As mentioned earlier, we consider four parallel processes on one side of the communication link. Figure 3 shows the system at one side of the link: there is a source process which generates messages, a sink process which consumes messages that come from the opposite side of the communication link, and the transmission station, which consists of one monitor which is called upon by the source and sink processes, and two processes that look after the transmission and reception of frames. The processes execute certain procedures which refer to the central monitor or to the transmission line, as indicated by the arrows. The sender and receiver processes are synchronized with the speed of the transmission line. Relative to this speed, the source and sink processes are synchronized by the control monitor.

Figure 3: The structure of the communication system at one side of the communication line.

The inner structure of the control monitor, not shown in Figure 3, is described by the Concurrent Pascal 11 program of the appendix. Essentially the monitor consists of the four procedures called upon by the processes. It enforces the mutual exclusion between the execution of these procedures. It contains the sequence counters L, A, S, and R, and a finite size send-buffer. Two queue variables are used to make the source or sink process wait, if necessary.

We note that very few logical changes have been introduced to the protocol definitions of the section 3.1 and 3.3 in order to obtain this program. We conclude that the language used is suitable for programming communication protocols. We hope that, in the future, sufficiently efficient implementations of languages for parallel processes and monitors will be available for the implementation of protocols.

We note that in Figure 3, we have assumed the existence of two additional processes, not shown in Figure 3 and may be implemented in hardware, which actually perform the continuous transmission and reception of bit strings on the two lines. We call these processes line processes. The procedures transmit (Frame) and listen (Frame) of the sender and receiver processes refer to the actions of these line processes. The line processes should not be interrupted by the sender or receiver processes when the latter executes a control monitor procedure. The interaction between the sender, receiver, and line processes could again be described by using monitors. However, we do not intend to give a description of the line processes in this paper.

5. Conclusions

We believe that the design, verification, and implementation of a protocol are inter-related activities. In this article, we have considered a particular protocol as an example, and have shown how certain tools, developed in different contexts, can be successfully applied to the logical verification and implementation of this protocol. We are confident that the same methods can also be applied in the case of the more complex protocols that are used in actual applications. The main objective of these efforts is to obtain more reliable and better documented communication systems.

We note that, apart from the tools discussed in this paper, other approaches (see for example in reference 3) have been described for dealing with the problems of logical verification and efficient implementation of communication protocols. Further research is needed for comparing the relative merits of these different methods.

There are at least two remaining problems which, in this paper, have been mentioned only briefly. The first problem (see section 3.4) is related to the
logical verification of protocols, and is expressed in
the following questions: What is the best method for
proving that the operation of a given protocol, for a
finite amount of information transfer, will terminate
after a finite amount of time? What is the best method
for showing that there are no undesired loops or dead-
locks in the communication system?

The other problem is related to the implementation
of protocols. We have proposed the use of a well-
structured high-level language for describing parallel
processes and monitors. This approach supposes that an
efficient implementation of such a language exists on
the computers that are used for the communication sys-
tem. The author does not know of any language imple-
mentation of this kind that is available at present.
However, several projects of implementation are in
progress.

Acknowledgements:
I am grateful to Jean Vaucher for his interesting
discussions, and many useful suggestions about the
content of this paper.

APPENDIX: The control monitor of a transmission sta-
tion written in Concurrent Pascal.

const
mseq = 2 * M;
shuf = ... {multiple of mseq};
type
sequence-count = 0 .. shuf-1;
buffher-index = 0 .. shuf-1;
message-type = ...;
frame-type = record S-frame, R-frame :
msg-frame: message-type;
check-field: ...
end;
var
control-monitor: monitor;
var send-buffer: array [buffer-index] of
message-type;
A, L: buffer-index;
S, R: sequence-count;
next-frame, buffer-free: queue;
sink-waiting: boolean;
sink-pointer: message-type;
end;

begin {monitor initialization}
A := L := S := R := 0;
sink-waiting := false;
clear (next-frame); clear (buffer-free)
end;

REFERENCES:
[1] L. POUZIN, Network protocols, Nato International
[3] ACM Interprocess Communication Workshop, Santa
Monica, Calif., March 1975.
[4] G.V. KOCHMANN, Communication protocols and error
recovery procedures, in reference 3.
[5] P. BRINCH-HANSEN, Operating systems principles,
[6] C.A.R. HOARE, Monitors: An operating system struc-
turing concept, Communications ACM, 17, p. 549
(1974).
control: a perspective, IBM Systems Journal, 13,
[9] DATAPAC, Standard network access protocol, The
computer communication group, Trans-Canada Tele-
[10] L. POUZIN, Basic elements of a network data link
control system, AOM Computer Communication
tors and condition variables, in reference 3.
[12] P. BRINCH-HANSEN, Concurrent Pascal - a program-
ming language for operating systems design, Tech.
Report No. 11, Information Science, Cal Tech,
April 1974.