INTERNATIONAL TELECOMMUNICATION UNION

CCITT

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

H.261

(11/1988)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

Coding of moving video

CODEC FOR AUDIOVISUAL SERVICES AT n × 384 kbit/s

Reedition of CCITT Recommendation H.261 published in the Blue Book, Fascicle III.6 (1988)

NOTES

1 CCITT Recommendation H.261 was published in Fascicle III.6 of the *Blue Book*. This file is an extract from the *Blue Book.* While the presentation and layout of the text might be slightly different from the *Blue Book* version, the contents of the file are identical to the *Blue Book* version and copyright conditions remain unchanged (see below).

2 In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Recommendation H.261

CODEC FOR AUDIOVISUAL SERVICES AT n × 384 kbit/s

(Melbourne, 1988)

The CCITT,

considering

(a) that there is significant customer demand for videoconference service;

(b) that circuits to meet this demand can be provided by digital transmission using the H_0 rate or its multiples up to the primary rate;

(c) that ISDNs are likely to be available in some countries that provide a switched transmission service at the H_0 rate;

(d) that the existence of different digital hierarchies and different television standards in different parts of the world complicates the problems of specifying coding and transmission standards for international connections;

(e) that videophone services are likely to appear using basic ISDN access and that some means of interconnection of videophone and videoconference terminals should be possible;

(f) that Recommendation H.120 for videoconferencing using primary digital group transmission was the first in an evolving series of Recommendations,

appreciating

that advances are being made in research and development of video coding and bit rate reduction techniques which will lead to further Recommendations for videophone and videoconferencing at multiples of 64 kbit/s during subsequent Study Periods, so that this may be considered as the second in the evolving series of Recommendations.

and noting

that it is the basic objective of CCITT to recommend unique solutions for international connections,

recommends

that in addition to those codecs complying with Recommendation H.120, codecs having signal processing and interface characteristics described below should be used for international videoconference connections.

Note 1 – Codecs of this type are also suitable for some television services where full broadcast quality is not required.

Note 2 – Equipment for transcoding from and to codecs according to Recommendation H.120 is under study.

Note 3 – It is recognised that the objective is to provide interworking between $n \times 384$ kbit/s codecs and $m \times 64$ kbit/s codecs as defined in the H-Series Recommendations. Interworking will be on the basis of $m \times 64$ kbit/s, where the values of m are under study.

1 Scope

This Recommendation describes the coding and decoding methods for audiovisual services at the rates of $n \times 384$ kbit/s, where n is 1 to 5. Possible extension of this scope to meet the objective in Note 3 above is under study.

2 Brief specification

An outline block diagram of the codec is given in Figure 1/H.261.

2.1 *Video input and output*

To permit a single Recommendation to cover use in and between 625 and 525-line regions, pictures are coded in one common intermediate format. The standards of the input and output television signals, which may, for example, be composite or component, analogue or digital and the methods of performing any necessary conversion to and from the intermediate coding format are not subject to recommendation.

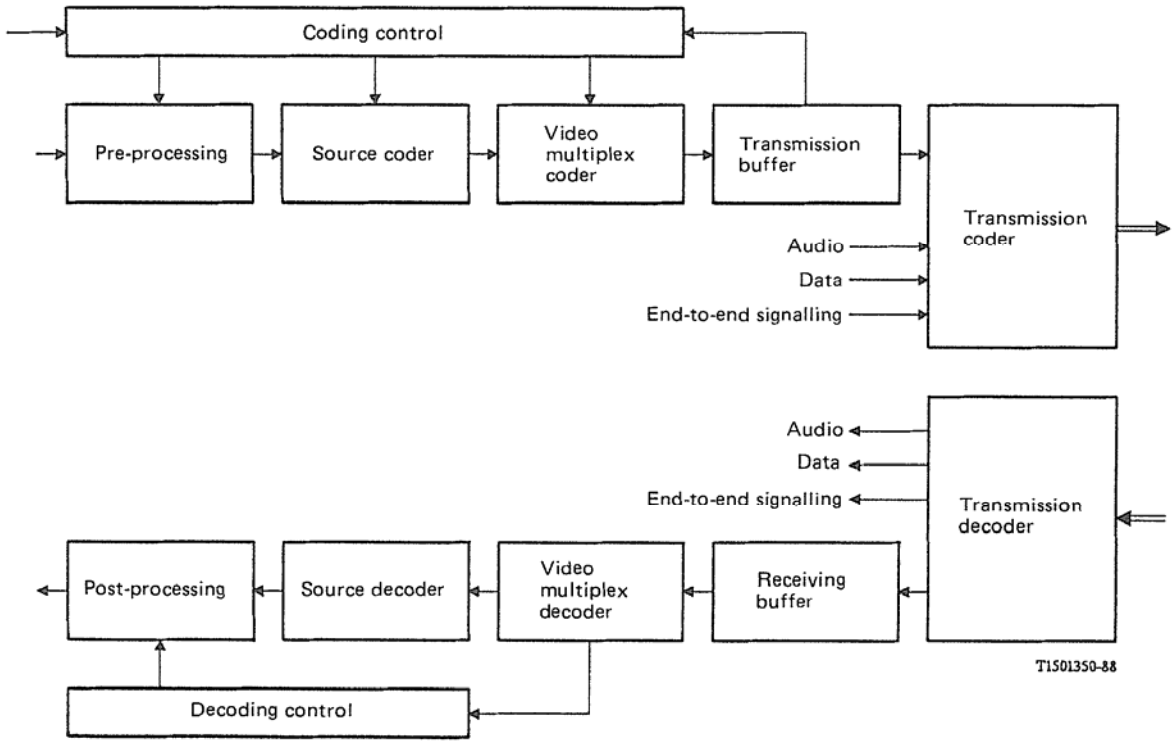


FIGURE 1/H.261

Outline block diagram of the codec

2.2 *Digital output and input*

Digital access at the primary rate of 1544 or 2048 kbit/s is with vacated time slots in accordance with Recommendation I.431.

Interfaces using ISDN basic accesses are under study (see Recommendation I.420).

2.3 *Sampling frequency*

Pictures are sampled at an integer multiple of the video line rate. This sampling clock and the digital network clock are asynchronous.

2.4 *Source coding algorithm*

A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is adopted. The decoder has motion compensation capability, allowing optional incorporation of this technique in the coder.

2.5 *Audio channel*

Audio is coded according to mode 2 of Recommendation G.722. This is combined with control and indication information and conveyed in one 64 kbit/s time slot which conforms to Recommendation H.221.

2.6 *Data channels*

Recommendation H.221 permits part of the 64 kbit/s time slot carrying the audio to be used for auxiliary data transmission.

Additionally, one of the time slots normally used for video may be reassigned as a 64 kbit/s data channel. The possibility of further such channels is under study.

2.7 *Symmetry of transmission*

The codec may be used for bidirectional or unidirectional audiovisual communication.

2.8 *Error handling*

Under study.

2.9 *Propagation delay*

Under study.

2.10 *Additional facilities*

Under study.

3 Source coder

3.1 *Source format*

The source coder operates on non-interlaced pictures occurring 30000/1001 (approximately 29.97) times per second. The tolerance on picture frequency is ± 50 ppm.

Pictures are coded as luminance and two colour difference components (Y, C_R et C_B). These components and the codes representing their sampled values are as defined in CCIR Recommendation 601.

Black = 16

White = 235

Zero colour difference = 128

Peak colour difference = 16 and 240.

These values are nominal ones and the coding algorithm functions with input values of 0 through to 255.

For coding, the luminance sampling structure is 288 lines per picture, 352 pels per line in an orthogonal arrangement. Sampling of each of the two colour difference components is at 144 lines, 176 pels per line, orthogonal. Colour difference samples are sited such that their block boundaries coincide with luminance block boundaries as shown in Figure 2/H.261. The picture area covered by these numbers of pels and lines has an aspect ratio of 4 : 3 and corresponds to the active portion of the local standard video input.

Note – The number of pels per line is compatible with sampling the active portions of the luminance and colour difference signals from 525 to 625-line sources at 6.75 and 3.375 MHz, respectively. These frequencies have a simple relationship to those in CCIR Recommendation 601.

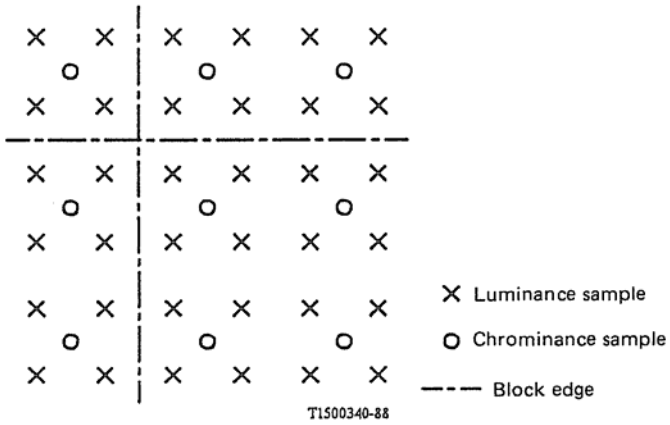
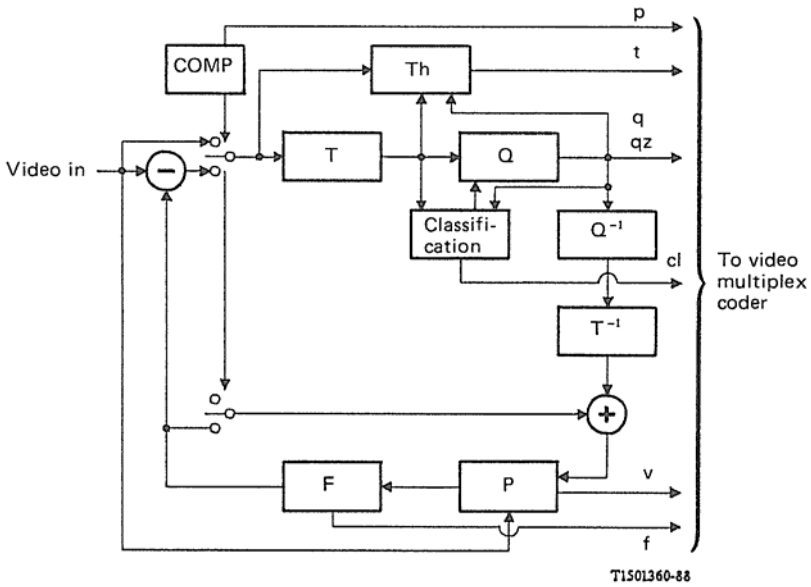


FIGURE 2/H.261

Positioning of luminance and chrominance samples

3.2 *Video source coding algorithm*

The video coding algorithm is shown in generalised form in Figure 3/H.261. The main elements are prediction, block transformation, quantization and classification.



COMP	Comparator for intra/inter
Th	Threshold
T	Transform
Q	Quantizer
P	Picture memory with motion compensated variable delay
F	Loop filter
p	Flag for intra/inter
t	Flag for transmitted or not
q	Quantizing index for transform coefficients
qz	Quantizer indication
v	Motion vector
cl	Classification index
f	Switching on/off of the loop filter

FIGURE 3/H.261

Video coding algorithm

The prediction error (INTER mode) or the input picture (INTRA mode) is subdivided into 8 pel by 8 line blocks which are segmented as transmitted or non-transmitted. The criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the data rate control strategy. Transmitted blocks are transformed and resulting coefficients are quantized and variable length coded.

3.2.1 *Prediction*

The prediction is inter-picture and may be augmented by motion compensation (§ 3.2.2) and a spatial filter (§ 3.2.3).

3.2.2 *Motion compensation*

Motion compensation is optional in the encoder. The decoder will accept one vector for each block of 8 pels by 8 lines. The range of permitted vectors is under study.

A positive value of the horizontal or vertical component of the motion vector signifies that the prediction is formed from pels in the previous picture which are spatially to the right or below the pels being predicted.

Motion vectors are restricted such that all pels referenced by them are within the coded picture area.

3.2.3 *Loop filter*

The prediction process may be modified by a two-dimensional spatial filter which operates on pels within a predicted block.

The filter is separable into one dimensional hironzontal and vertical functions. Both are non-recursive with coefficients of 1/4, 1/2, 1/4. At block edges, where one of the taps would fall outside the block, the peripheral pel is used for two taps. Full arithmetic precision is retained with rounding to 8 bit integer values at the 2-D filter output. Values whose fractional part is one half are rounded up.

The filter may be switched on or off on a block by block basis. The method of signalling is under study.

3.2.4 *Transformer*

Transmitted blocks are coded with a separable 2-dimensional discrete cosine transform of size 8 by 8. The input to the forward transform and output from the inverse transform have 9 bits. The arithmetic procedures for computing the transforms are under study.

Note – The output from the forward and input to the inverse are likely to be 12 bits.

3.2.5 *Quantization*

The number of quantizers, their characteristics and their assignment are under study.

3.2.6 *Clipping*

To prevent quantization distortion of transform coefficient amplitudes causing arithmetic overflow in the encoder and decoder loops, clipping functions are inserted. In addition to those in the inverse transform, a clipping function is applied at both encoder and decoder to the reconstructed picture which is formed by summing the prediction and the prediction error as modified by the coding process. This clipper operates on resulting pel values less than 0 or greater than 255, changing them to 0 and 255 respectively.

3.3 *Data rate control*

Sections where parameters which may be varied to control the rate of generation of coded video data include processing prior to the source coder, the quantizer, block significance criterion and temporal subsampling. The proportions of such measures in the overall control strategy are not subject to recommendation.

When invoked, temporal subsampling is performed by discarding complete pictures. Interpolated pictures are not placed in the picture memory.

3.4 *Forced updating*

This function is achieved by forcing the use of the INTRA mode of the coding algorithm. The update interval and pattern are under study.

4 Video multiplex coder

4.1 *Data structure*

Note 1 – Unless specified otherwise, the most significant bit is transmitted first.

Note 2 – Unless specified otherwise, bit 1 is transmitted first.

Note 3 – Unless specified otherwise, all unused or spare bits are set to ′1′.

4.2 *Video multiplex arrangement*

4.2.1 *Picture header*

The structure of the picture header is shown in Figure 4/H.261. Picture headers for dropped pictures are not transmitted.

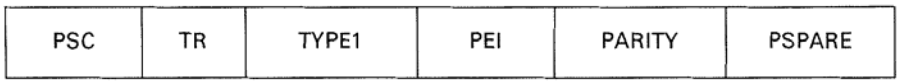


FIGURE 4/H.261

Structure of picture header

4.2.1.1 *Picture start code (PSC)*

A unique word of 21 bits which cannot be emulated by error-free data. Its value is under study.

4.2.1.2 *Temporal reference (TR)*

A five bit number derived using modulo-32 counting of pictures at 29.97 Hz.

4.2.1.3 *Type information (TYPE1)*

Information about the complete picture:

Bit 1	Split screen indicator. ′0′ off, ′1′ on.
Bit 2	Document camera. ′0′ off, ′1′ on.
Bit 3	Freeze picture release. Under study.
Bit 4	Under study. Possible uses include signalling of the use of motion compensation and the method of switching the loop filter.
Bit 5	Number of classes. ′0′ one, ′1′ four.
Bits 6 to 12	Under study.

4.2.1.4 *Extra insertion information (PEI)*

Two bits which signal the presence of the following two optional data fields.

4.2.1.5 *Parity information (PARITY)*

For optional use and present only if the first PEI bit is set to ′1′. Eight parity bits each representing odd parity of the aggregate of the corresponding bit planes of the locally decoded PCM values of Y, C_R and C_B in the previous picture period.

4.2.1.6 *Spare information (PSPARE)*

Sixteen bits are present when the second PEI bit is set to '1'. The use of these bits is under study.

4.2.2 *Group of blocks header*

A group of blocks consists of 2k lines of 44 luminance blocks each, k lines of 22 C_R blocks and k lines of 22 C_B blocks. The value of k is under study.

The structure of the group of blocks header is shown in Figure 5/H.261. All GOB headers are transmitted except those in dropped pictures.

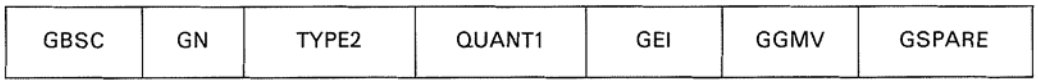


FIGURE 5/H.261

Structure of group of blocks header

4.2.2.1 *Group of blocks start code (GBSC)*

A word of 16 bits, 0000 0000 0000 0001.

4.2.2.2 *Group number (GN)*

An *m* bit number indicating the vertical position of the group of blocks. The value of *m* is the smallest integer greater than or equal to $\log_2$ (18/k). GN is 1 at the top of the picture.

Note – GBSC plus the following GN is not emulated by error-free video data.

4.2.2.3 *Type information (TYPE2)*

TYPE2 is *p* bits which give information about all the transmitted blocks in a group of blocks. The value of *p* is under study.

Bit 1 When set to '1' indicates that all the transmitted blocks in the GOB are coded in INTRA mode and without block addressing data.

Bits 2 to *p* Spare, under study.

4.2.2.4 *Quantizer information (QUANT1)*

A *j* bit code word which indicates the blocks in the group of blocks where QUANT2 code words are present. These blocks, their code words and the value of *j* are under study.

Whether QUANT1 is in the GOB header or the picture header is under study.

4.2.2.5 *Extra insertion information (GEI)*

Under study.

4.2.2.6 *Group of blocks global motion vector (GGMV)*

Under study.

4.2.2.7 *Spare information (GSPARE)*

Under study.

4.2.3 *Block data alignment*

The structure of the data for *n* transmitted blocks is shown in Figure 6/H.261. The values of *n* and the order are under study. Elements are omitted when not required.

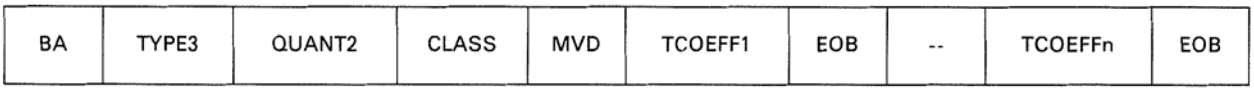


FIGURE 6/H.261

Data structure of transmitted block

4.2.3.1 *Block address (BA)*

A variable length code word indicating the position of *n* blocks within a group of blocks. VLC code words using a combination of relative and absolute addressing are under study.

The transmission order and addressing of blocks are under study.

When bit 1 of TYPE2 is '1', BA is not included and up to 132k blocks beginning with and continuing in the above transmission order are transmitted before the next GOB header.

4.2.3.2 *Block type information (TYPE3)*

Variable length code words indicating the types of blocks and which data elements are present. Block types and VLC code words are under study.

4.2.3.3 *Quantizer (QUANT2)*

A code word of up to q bits signifying the table(s) used to quantize transform coefficients. The value of q and the code words are under study. QUANT2 is present in the first transmitted block after the position indicated by QUANT1.

4.2.3.4 *Classification index (CLASS)*

CLASS is present if bit 5 of TYPE1 is set to '1' and indicates which of the four available transmission sequence orders is used for luminance block coefficients. If bit 5 of TYPE1 is set to '0' then luminance block coefficients are transmitted in the default sequence order.

Chrominance block coefficients are transmitted in one sequence order.

The CLASS code words and sequence orders are under study.

4.2.3.5 *Motion vector data (MVD)*

Calculation of the vector data is under study.

When the vector data is zero, this is signalled by TYPE3 and MVD is not present.

When the vector data is non-zero, MVD is present consisting of a variable length code word for the horizontal component followed by a variable length code word for the vertical component.

Variable length coding of the vector components is under study.

4.2.3.6 *Transform coefficients (TCOEFF)*

The quantized transform coefficients are sequentially transmitted according to the sequence defined by CLASS. The DC component is always first. Coefficients after the last non-zero one are not transmitted.

The coding method and tables are under study.

4.2.3.7 *End of block marker (EOB)*

Use of and code word for EOB are under study. An EOB without any transform coefficients for a block is allowed.

4.3 *Multipoint considerations*

4.3.1 *Freeze picture request*

Causes the decoder to freeze its received picture until a picture freeze release signal is received. The transmission method for this control signal is under study.

4.3.2 *Fast update request*

Causes the encoder to empty its transmission buffer and encode its next picture in INTRA mode with coding parameters such as to avoid buffer overflow. The transmission method for this control signal is under study.

4.3.3 *Data continuity*

The prototocl adopted for ensuring continuity of data channels in a switched multipoint connection is handled by the message channel. Under study.

5 Vide data buffering

The size of the transmission buffer at the encoder and its relationship to the transmittion rate are under study.

Transmission buffer overflow and underflow are not permitted. Measures to prevent underflow are under study.

6 Transmission coder

6.1 *Bit rate*

The net bit rate including audio and optional data channels is an integer multiple of 384 kbit/s up to and including 1920 kbit/s.

The source and stability of the encoder output clock are under study.

6.2 *Video clock justification*

Video clock justification is not provided.

6.3 *Frame structure*

6.3.1 *Frame structure for 384-1920 kbit/s channels*

The frame structure is defined in Recommendation H.222.

6.3.2 *Bit assignment in application channel*

Under study.

6.3.3 *Time slot positioning*

According to Recommendation I.431.

6.4 *Audio coding*

Recommendation G.722 56/48 kbit/s audio, 0/8 kbit/s data and 8 kbit/s service channel in the first time slot.

The delay of the encoded audio relative to the encoded video at the channel output is under study.

6.5 *Data transmission*

One or more time slots may be allocated as data channels of 64 kbit/s each. The first channel uses the fourth time slot.

Positioning of the other channels, and possible restrictions on availability at lower overall bit rates are under study. The BAS codes used to signal that these data channels are in use are specified in Recommendation H.221.

6.6 *Error handling*

Under study.

6.7 *Encryption*

Under study.

6.8 *Bit sequence independence restrictions*

Under stydy.

6.9 *Network interface*

Access at the primary rate is with vacated time slots as per Recommendation I.431.

For 1544 kbit/s interfaces the default H_0 channel is time slots 1 to 6.

For 2048 kbit/s interfaces the default H_0 channel is time slots 1-2-3-17-18-19.

Interfaces using ISDN basic accesses are under study (see Recommendation I.420).

ITU-T H-SERIES RECOMMENDATIONS

AUDIOVISUAL AND MULTIMEDIA SYSTEMS

Characteristics of transmission channels used for other than telephone purposes	H.10–H.19
Use of telephone-type circuits for voice-frequency telegraphy	H.20–H.29
Telephone circuits or cables used for various types of telegraph transmission or simultaneous transmission	H.30–H.39
Telephone-type circuits used for facsimile telegraphy	H.40–H.49
Characteristics of data signals	H.50–H.99
CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	H.220–H.229
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	**H.260–H.279**
Related systems aspects	H.280–H.299
Systems and terminal equipment for audiovisual services	H.300–H.399
Supplementary services for multimedia	H.450–H.499

For further details, please refer to ITU-T List of Recommendations.

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	**Audiovisual and multimedia systems**
Series I	Integrated services digital network
Series J	Transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland

Geneva, 2006