課題　データ圧縮 (ハフマン符号化)

ハフマン符号化に基づくデータ圧縮と解凍プログラムを作成し，次ページのデータをファイルに保存し、圧縮前と圧縮後でのファイルサイズを比較せよ．また，圧縮されたデータから元のデータを復元できることを確認せよ．

ハフマン符号化

ハフマン符号化は，文字の出現頻度に応じてその文字を表現するビット長を変えることでデータを圧縮する．出現頻度の高い文字は短い符号長で表現し，出現頻度の低いものは長い符号長で表現する．具体的な方法を以下に示す．いま例として文章中に A, B, C, D, E, F, G, H の 7 文字のみが現れる場合を考える．これを 3bit 固定長で符号化しようとすると例えば以下のように符号を割り当てる．

文字	A	B	C	D	E	F	G	H
符号	000	001	010	011	100	101	110	111

通常 1 文字は 8 ビットで表現されるので，このときデータは 3/8 = 37.5%に圧縮されることになる．しかしながら，当然ながら文章中に現れる文字の種類が増えると圧縮率は低下する．

ハフマン符号化では各文字の出現頻度に応じて以下のように符号化する．

文字	A	B	C	D	E	F	G	H
出現頻度	30%	24%	20%	16%	6%	2%	1%	1%
符号	00	10	11	010	0110	01110	011110	011111

このとき 1 文字の平均符号長は 2.42 bit であり，圧縮率は 2.42/8 = 30.3%である．

ハフマン符号化の具体的な作り方は，以下の通りである．出現頻度の最も低い 2 つをひとつにまとめ，そのどちらかの出現頻度をその和として考える．以下同様に，この手順を繰り返していく．

A	30%
B	24%
C	20%
D	16%
E	6%
F	2%
G	2%
H	1%

A	30%
B	24%
C	20%
D	16%
E	6%
F	2%
(G,H)	2%

A	30%
B	24%
C	20%
D	16%
E	6%
(F,(G,H)	4%

A	30%
B	24%
C	20%
D	16%
(E,(F,(G,H))	10%

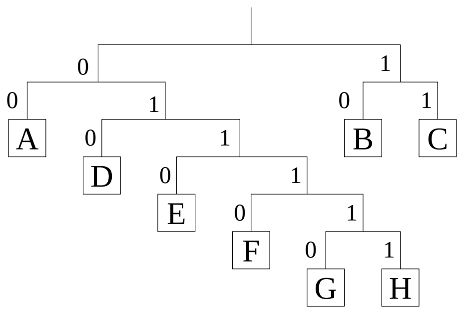
A	30%
(D,(E,(F,(G,H)))	26%
B	24%
C	20%

(B,C)	44%
A	30%
(D,(E,(F,(G,H)))	26%

(A,(D,(E,(F,(G,H))))	56%
(B,C)	44%

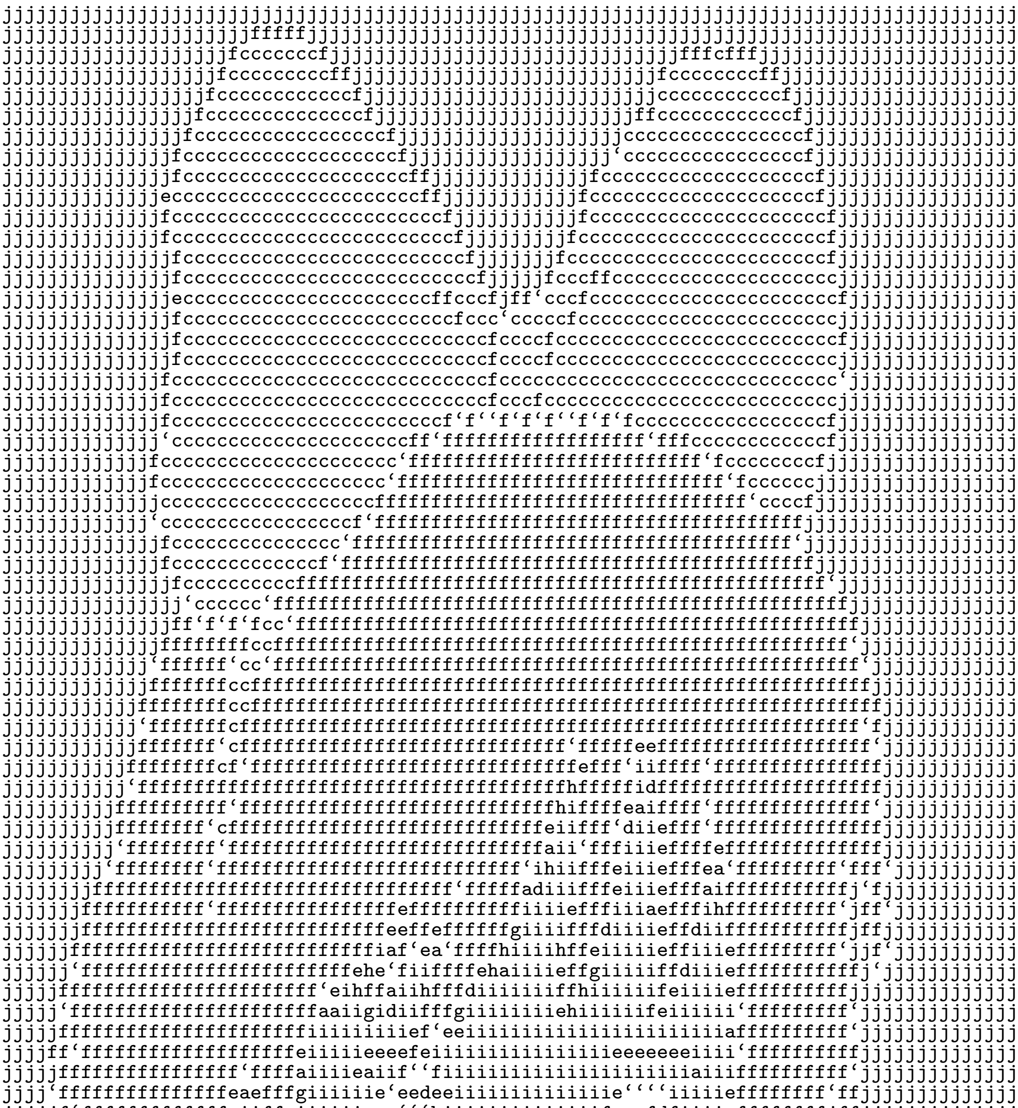
((A,(D,(E,(F,(G,H)))),(B,C))	100%

これを木構造で表現すると以下のようになる．左の枝に 0，右の枝に 1 を割り当て，各文字に到達するまで通った枝の数値を読み上げるとそれがその文字に対応する符号になる．



(実行例)

2 ページに示すデータを圧縮した結果，ファイルサイズは 2824/8190 = 34.5 % に縮小された．
3 ページに示す (72 種類の文字が含まれる) 光ファイバに関するテキストデータを圧縮した結果，ファイルサイズは 5537/7987 = 69.3%に縮小された．



Optical fiber

An optical fiber is made up of the core(carrying the light pulses), the cladding (reflecting the light pulses back into the core) and the buffer coating (protecting the core and cladding from moisture, damage etc.). Together, all of this creates a fiber optic which can carry up to 10 million messages at any time using light pulses. Fiber optics is the overlap of applied science and engineering concerned with the design and application of optical fibers. Optical fibers are widely used in fiber-optic communications, which permits transmission over longer distances and at higher bandwidths (data rates) than other forms of communications. Fibers are used instead of metal wires because signals travel along them with less loss, and they are also immune to electromagnetic interference. Fibers are also used for illumination, and are wrapped in bundles so they can be used to carry images, thus allowing viewing in tight spaces. Specially designed fibers are used for a variety of other applications, including sensors and fiber lasers.

Light is kept in the core of the optical fiber by total internal reflection. This causes the fiber to act as a waveguide. Fibers which support many propagation paths or transverse modes are called multi-mode fibers (MMF), while those which can only support a single mode are called single-mode fibers (SMF). Multi-mode fibers generally have a larger core diameter, and are used for short-distance communication links and for applications where high power must be transmitted. Single-mode fibers are used for most communication links longer than 550 meters (1,800 ft).

Joining lengths of optical fiber is more complex than joining electrical wire or cable. The ends of the fibers must be carefully cleaved, and then spliced together either mechanically or by fusing them together with an electric arc. Special connectors are used to make removable connections.

Fiber optics, though used extensively in the modern world, is a fairly simple and old technology. Guiding of light by refraction, the principle that makes fiber optics possible, was first demonstrated by Daniel Colladon and Jacques Babinet in Paris in the early 1840s. John Tyndall included a demonstration of it in his public lectures in London a dozen years later.[1] Tyndall also wrote about the property of total internal reflection in an introductory book about the nature of light in 1870: "When the light passes from air into water, the refracted ray is bent towards the perpendicular... When the ray passes from water to air it is bent from the perpendicular... If the angle which the ray in water encloses with the perpendicular to the surface be greater than 48 degrees, the ray will not quit the water at all: it will be totally reflected at the surface.... The angle which marks the limit where total reflection begins is called the limiting angle of the medium. For water this angle is 4827', for flint glass it is 3841', while for diamond it is 2342'."[2][3]

History

Practical applications, such as close internal illumination during dentistry, appeared early in the twentieth century. Image transmission through tubes was demonstrated independently by the radio experimenter Clarence Hansell and the television pioneer John Logie Baird in the 1920s. The principle was first used for internal medical examinations by Heinrich Lamm in the following decade. In 1952, physicist Narinder Singh Kapany conducted experiments that led to the invention of optical fiber. Modern optical fibers, where the glass fiber is coated with a transparent cladding to offer a more suitable refractive index, appeared later in the decade.[1] Development then focused on fiber bundles for image transmission. The first fiber optic semi-flexible gastroscope was patented by Basil Hirschowitz, C. Wilbur Peters, and Lawrence E. Curtiss, researchers at the University of Michigan, in 1956. In the process of developing the gastroscope, Curtiss produced the first glass-clad fibers; previous optical fibers had relied on air or impractical oils and waxes as the low-index cladding material. A variety of other image transmission applications soon followed.

The use of optic fibers for communication purposes were first carried out in Western Europe in the late 19th and early 20th century, such as they were used to diagnose a patient's stomach by a doctor, and those communications within short ranges. Especially, the transfer of images by optical fibers was largely popularized at the beginning of 21st century, due to the growing medical and television demands.[4]

It is said that, Jun-ichi Nishizawa, a Japanese scientist at Tohoku University, also proposed the use of optical fibers for communications, in 1963, as stated in his own book published in 2004 in India.[5] Nishizawa invented other technologies that contributed to the development of optical fiber communications as well.[6] Nishizawa later invented the graded-index optical fiber as a channel for transmitting light from semiconductor lasers.[7]

The groundbreaking event happened in around 1965, Charles K. Kao and George A. Hockham of the British company Standard Telephones and Cables (STC) were the first to promote the idea that the attenuation in optical fibers could be reduced below 20 decibels per kilometer (dB/km), allowing fibers to be a practical medium for communication.[8] They proposed that the attenuation in fibers available at the time was caused by impurities, which could be removed, rather than fundamental physical effects such as scattering. They correctly and systematically theorized the light-loss properties for optical fiber, and pointed out the right material to manufacture such fibers - silica glass with high purity. This discovery led to Kao being awarded the Nobel Prize in Physics in 2009.[9]

NASA used fiber optics in the television cameras that were sent to the moon. At the time its use in the cameras was 'classified confidential' and only those with the right security clearance or those accompanied by someone with the right security clearence were permitted to handle the cameras.[10]

The crucial attenuation level of 20 dB/km was first achieved in 1970, by researchers Robert D. Maurer, Donald Keck, Peter C. Schultz, and Frank Zimar working for American glass maker Corning Glass Works, now Corning Incorporated. They demonstrated a fiber with 17 dB/km attenuation by doping silica glass with titanium. A few years later they produced a fiber with only 4 dB/km attenuation using germanium dioxide as the core dopant. Such low attenuations ushered in optical fiber telecommunications and enabled the Internet. In 1981, General Electric produced fused quartz ingots that could be drawn into fiber optic strands 25 miles (40 km) long.[11]

Attenuations in modern optical cables are far less than those in electrical copper cables, leading to long-haul fiber connections with repeater distances of 70150 kilometers (4393 mi). The erbium-doped fiber amplifier, which reduced the cost of long-distance fiber systems by reducing or even in many cases eliminating the need for optical-electrical-optical repeaters, was co-developed by teams led by David N. Payne of the University of Southampton, and Emmanuel Desurvire at Bell Labs in 1986. The more robust optical fiber commonly used today utilizes glass for both core and sheath and is therefore less prone to aging processes. It was invented by Gerhard Bernsee in 1973 of Schott Glass in Germany.[12]

In 1991, the emerging field of photonic crystals led to the development of photonic-crystal fiber[13] which guides light by means of diffraction from a periodic structure, rather than total internal reflection. The first photonic crystal fibers became commercially available in 2000.[14] Photonic crystal fibers can be designed to carry higher power than conventional fiber, and their wavelength dependent properties can be manipulated to improve their performance in certain applications.