Data corruption

Data corruption is the term used by computer personnel to explain how and why a computer sporadically produces incorrect or unreadable data. It is not just the concern of these experts, however: it should concern everybody from the person whose name unexpectedly appears on a bad debt register to the general practitioner with his surgery computer. The main causes of data corruption are defects in the hardware (for example, in the integrated circuit chips) and problems concerning the user and the software—the programs.

Three parts of the computer hardware may cause data corruption: the integrated circuits, the magnetic storage media (tapes or discs), and power supplies.

Modern integrated circuits are very reliable and have lifetimes of more than 10 years, but they are easily damaged by inappropriate handling. In practice defects in the circuits are rarely a problem; much more frequently data corruption results from damage to the storage media. For example, a microscopic speck of grit may damage the surface of a disc and erase information or make it unreadable. A minute scratch has a large effect because it disorganises the data sequence. All large computer installations filter and humidify incoming air—and staff working in the computer area should be forbidden to smoke.

Any part of the computer system may malfunction if the power supply unit does not remove transient high voltages in the mains. Ideally a power unit should also contain an auxiliary battery so that if the mains fails there is enough power for the computer to close down its operation in an orderly way.

Vandals may gain access to the computer by a telephone line or by using an unguarded keyboard. Computer vandals usually have considerable skill and may find a route to the deepest software recesses of the computer and there cause extensive corruption.

An inexpert computer user may accidentally enter data that are not displayed on the screen but which form characters in the memory. These are then transmitted to the storage media and stay there until they cause trouble later on. Well designed software should trap such characters and notify the user. Unfortunately, inexpert users have a knack of doing the most unexpected things and somehow even the best traps may fail.

Nevertheless, the main source of data corruption is poorly designed software. Given that information is typed into the computer correctly, a program should check the entry, where it is to be put, and how it is to be arranged and used. The problem is that when a software program has many options it may be impossible for the designer to test them all. For example, in a program where information from the user causes the flow to branch five times there are 32 possible paths to completion. In a larger program (such as that used by a general practitioner to store data on patients) the number of possible paths is enormous, and few software firms could guarantee to have checked all the permutations.

What, then, can be done to prevent data corruption and to cure it once present? To deal with the hardware first, the preventive approach is based on providing the computer with an optimum environment and a protected mains supply with auxiliary battery facilities. At the least a "No Smoking" room should be set aside for the computer, where all its media and supplies can be stored securely.

Data corruption caused by software can be prevented only by good design. For this reason anyone buying software should talk to unbiased users to see if they have had any problems with the proposed software. The conservative approach is to buy software that has been available for a year or two—since all its errors are likely to have been detected and corrected.

A J ASBURY

Senior Lecturer in Anaesthesia and Honorary Consultant,
Western Infirmary,
Glasgow G11 6NT

Regular Review

Mammography and screening for breast cancer

E J ROEBUCK

Almost 30 years ago Gershon-Cohen and his coworkers were the first to suggest that mass *x* ray surveys might be used in the early detection of breast cancer and showed that serial mammography at six month intervals could detect small, presumably early, breast cancers.[1,2] They found a lower incidence of metastatic spread to the lymph nodes in patients whose cancer was first discovered in this way.

The principal impetus for the development of screening for breast cancer came from the Hospital Insurance Plan study in New York inaugurated in 1963.[3] This randomised controlled trial, using mammography and palpation, was the first to show a reduction in mortality (which has been maintained at 30% over a period of 10 to 14 years[4]). The early results of the New York study, published in 1973, stimulated further studies designed to answer the operational questions which would follow the introduction of breast cancer screening, and the American breast cancer demonstration (BCDD) project[5] and the west London screening programme[6] convincingly showed that mammography had a higher specificity and sensitivity than did clinical examination.