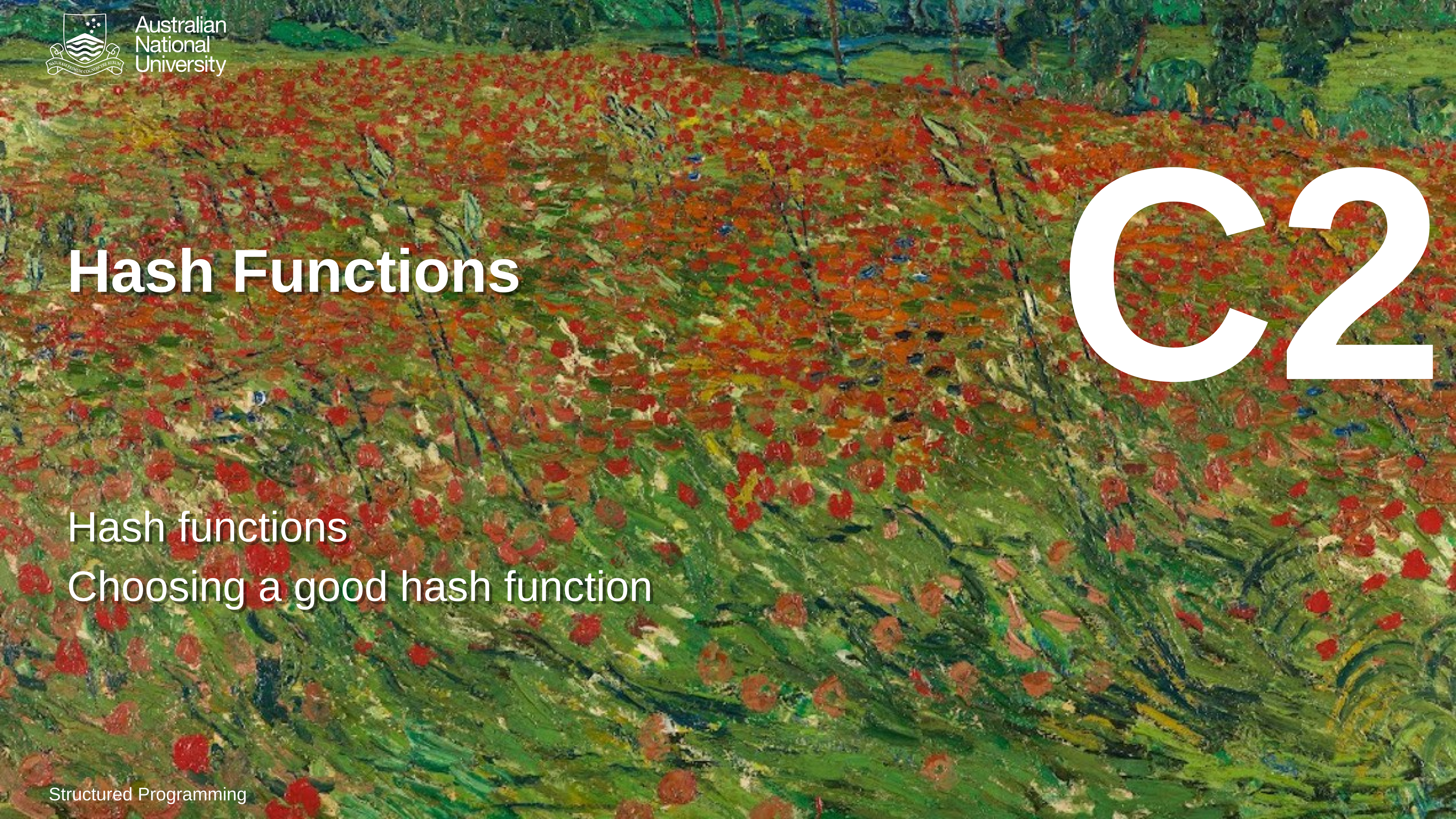


Australian National University

C2

Hash Functions

Hash functions

Choosing a good hash function

Hash Functions

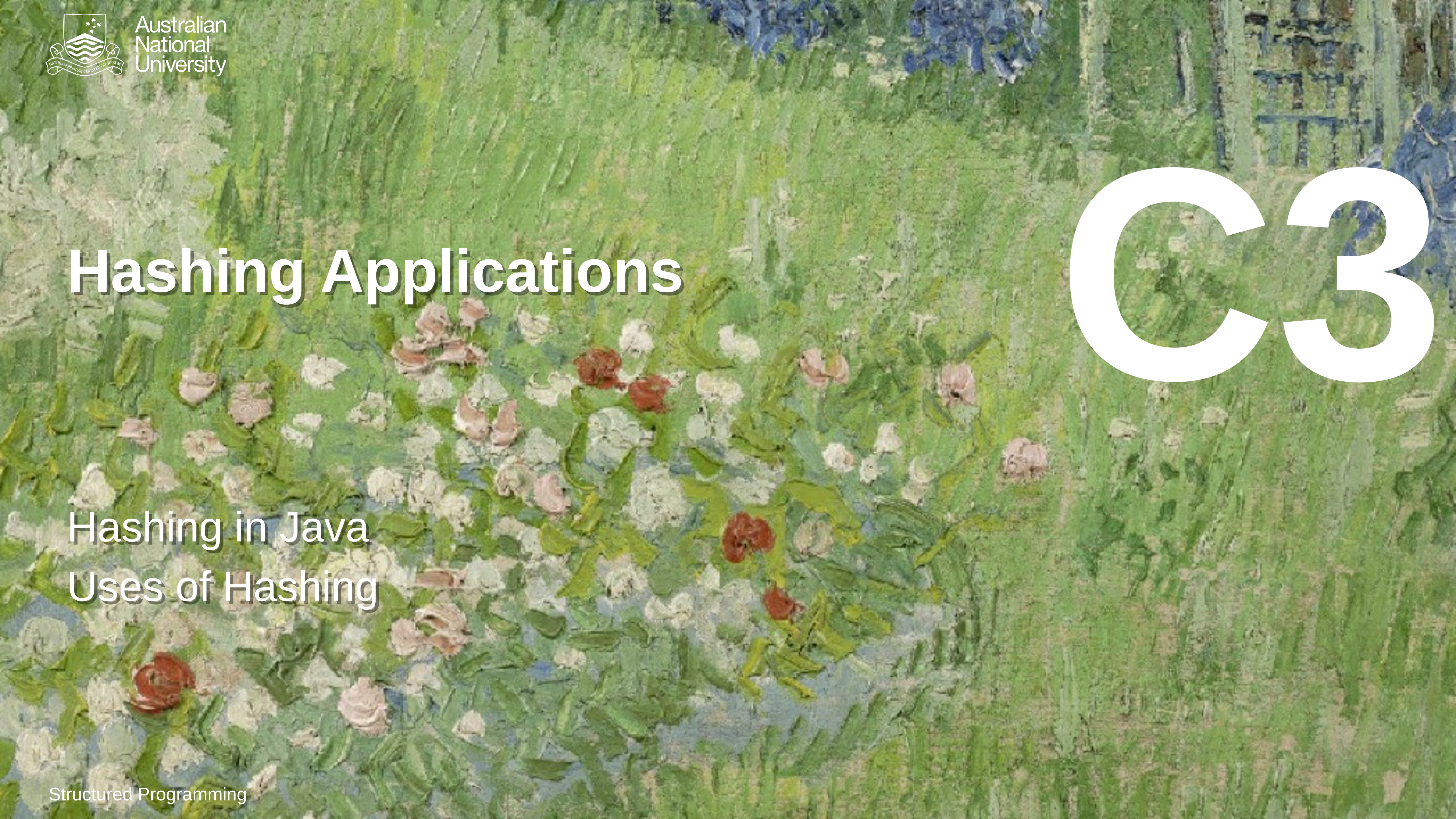
A hash function is a function *f(k)* that maps a key, *k*, to a value, *f(k)*, within a prescribed range.

A hash function is deterministic: for a given key *k*, *f(k)* is always the same (at least within a run of the program).

Choosing a Good Hash Function

A good hash for a given population, P, of keys, $k \in P$, will distribute $f(k)$ evenly within the prescribed range for the hash.

A *perfect hash* will give a unique $f(k)$ for each $k \in P$



Australian
National
University

Hashing Applications

C3

Hashing in Java

Uses of Hashing

Uses of Hashing

- Hash table (set or map implementation)
- Checksums
  - Error detection and/or correction.
- Compression
  - A hash is typically much more compact that the key
- Pruning a search
  - Looking for duplicates
- Cryptographic

Hashing Applications C3

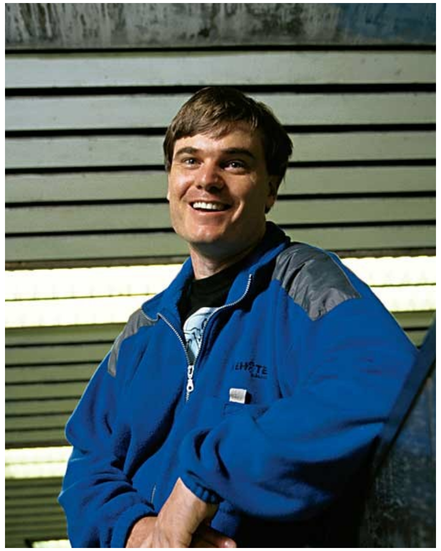
Luhn Algorithm

Used to check for transcription errors in credit cards (last digit checksum).

Hamming Codes

Error correcting codes (as used in EEC memory)

Hashing Applications C3



rsync (Tridgell)

Synchronize files by (almost) only moving the parts that are different.

MD5 (Rivest)

Previously used to encode passwords (but no longer).