Hashes

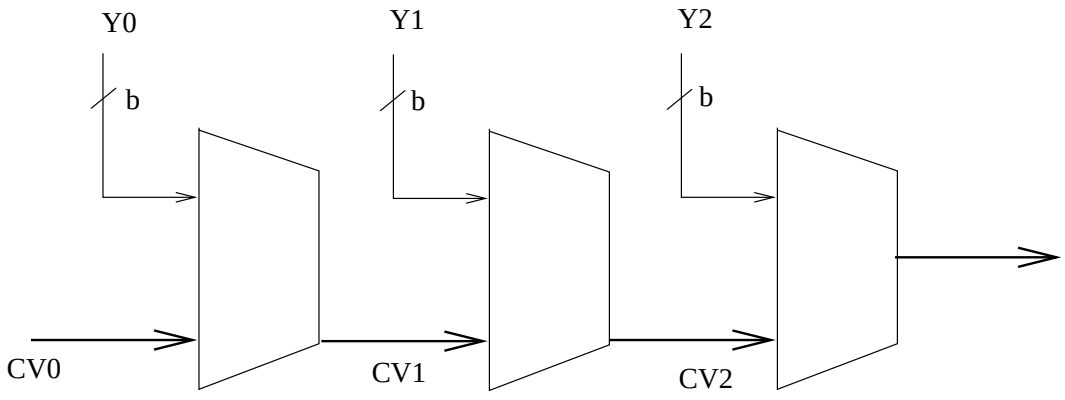
A hash function accepts a variable-size message m as input and produces a fixed-size hash code $h(m)$, called its message digest. It is a function of all the bits of the message. It should be[Sta99, Section 8.4]

- Relatively easy to compute.
- Pre-image resistant. Means a one-way hash i.e., given $y = h(x)$, can't find x.
- Second pre-image resistant. Useful for virus protection. Given x, can't find $x' \mid h(x) = h(x')$.
- Collision resistant. Can't find $x, y \mid h(x) = h(y)$ by just examining h.

A simple hash function is the XOR of fixed sized message blocks. Useless for data security. Trivial to compute pre-image and second pre-image.

By the birthday paradox, if the hash size is 64 bits, then time for collision $\approx 2^{32}$ (small). Typical hash size ≈ 160 bits.

Merkle-Damgård iterated construction: Can get collision resistance for arbitrary length strings from fixed length strings. Padding is 10^*, length is 64 bits $= 2^{31}$ GB. To put it another way, to construct a CRHF, it suffices to construct a collision-resistant compression function.



SHA1

Show [Sta99, Fig. 9.5, 9.6].

Customized Hash Functions

Name	Speed	Hash Length	Comments
MD4		128 bits	Proprietary. $O(2^{26})$ for inversion! Not even one-way.
MD5	28.5MB/s	128 bits	On ≈ 200 MHz Pentium. Considered weak now.
SHA-1	15.2MB/s	160 bits	NIST
RIPE-MD	12.6MB/s	160 bits	

These are much faster than block ciphers and heavily used. The future is AHA (Advanced Hashing Algorithm) with variable bit output — 160, 192, 256 bits.

References

[BCK96a] Mihir Bellare, Ran Canetti, and Hugo Krawczyk. Keying hash functions for message authentication. *Proceedings of Crypto*, 1996. An expanded version is available at http://www-cse.ucsd.edu/users/mihir.

[BCK96b] Mihir Bellare, Ran Canetti, and Hugo Krawczyk. Message authentication using hash functions—the hmac construction. Technical report, RSA Laboratories, 1996.

[Sta99] William Stallings. *Cryptography and Network Security*. Prentice-Hall Inc., 2nd edition, 1999.