Hash functions

- A hash function is a function whose output is shorter than its input.
- SHA1: \( \{0,1\}^{<2^{64}} \rightarrow \{0,1\}^{160} \)
- Standardized by NIST.
- Design principles are similar to that of other hash functions MD4 and MD5 proposed by Rivest.
- The inputs are first padded and divided by blocks. Then an iterated (chaining) compression function is applied (known as Merkle-Damgård transform):

![Diagram of hash function]

Security of hash functions

- What security properties a good hash function \( H \) should have?
- Collision-resistance (very informally): nobody should be able to efficiently find \( M_1, M_2 \) s.t. \( H(M_1) = H(M_2) \)
- One-wayness (very informally): nobody given \( h \) should be able to find \( M \) s.t. \( H(M) = h \).

Looking for collisions

- Let’s recall (learn) the “birthday” paradox.
- Applying the birthday-attack strategy and some additional analysis one can see that after making \( q \approx \sqrt{2^N} \) hash computations one can find collisions with probability close to 1
- Here \( N \) is the size of the range.
- So for SHA1 approximately \( 2^{80} \) trials will suffice.
Are more efficient attacks possible?

• Yes. Very recently collisions were found for MD4, MD5.

• February 2005. Xiaoyun Wang, Lisa Yiqun Yin, and Hongbo Yu described the way to find collisions in SHA1 by using $2^{69}$ hash computations (much faster than the birthday attack).

• February 2005. The result by Xiaoyun Wang, Andrew Yao and Frances Yao is announced. Collisions in SHA1 can be found by using $2^{63}$ hash computations.

• The attacks were not implemented and still does not appear very practical.

• But the standard SHA1 will most probably be replaced.