Hamiltonian Cycle problem

——From game to DNA computer

Dongbo Bu

Institute of Computing Technology

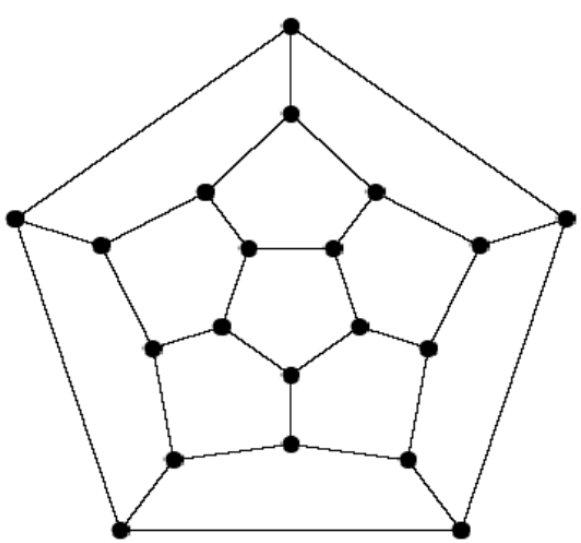
Chinese Academy of Sciences

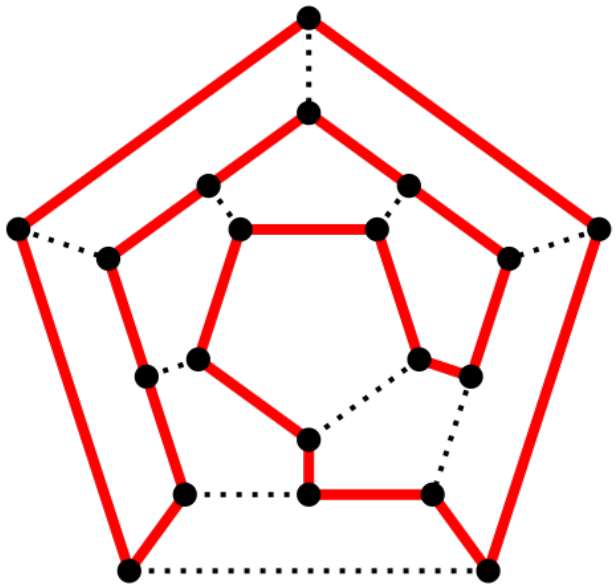
2016 11 29

Hamilton Cycle Problem

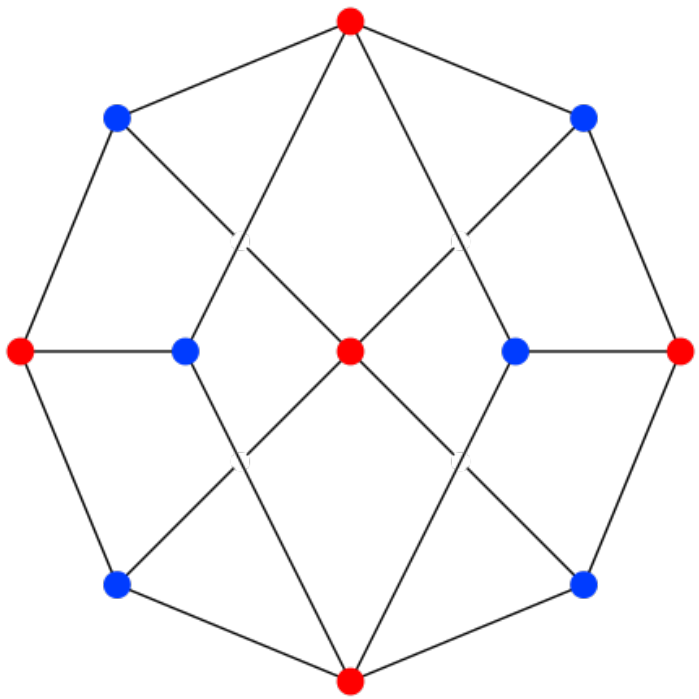
- In 1857, William Hamilton invented a game to find a cycle in an edge graph of dodecahedron.

A graph that has a Hamilton Cycle

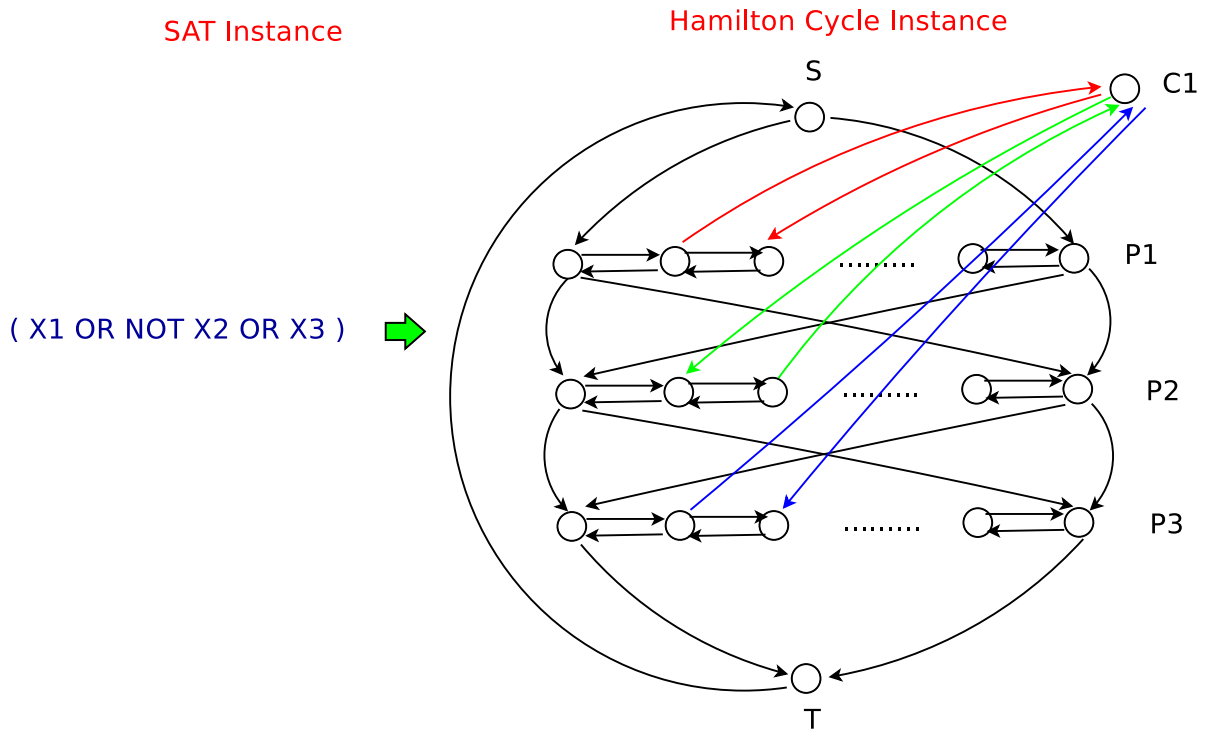




A graph that does not have a Hamilton Cycle



It is a hard problem to determine whether a graph has a $3\text{SAT} \leq_P$ Hamilton Cycle

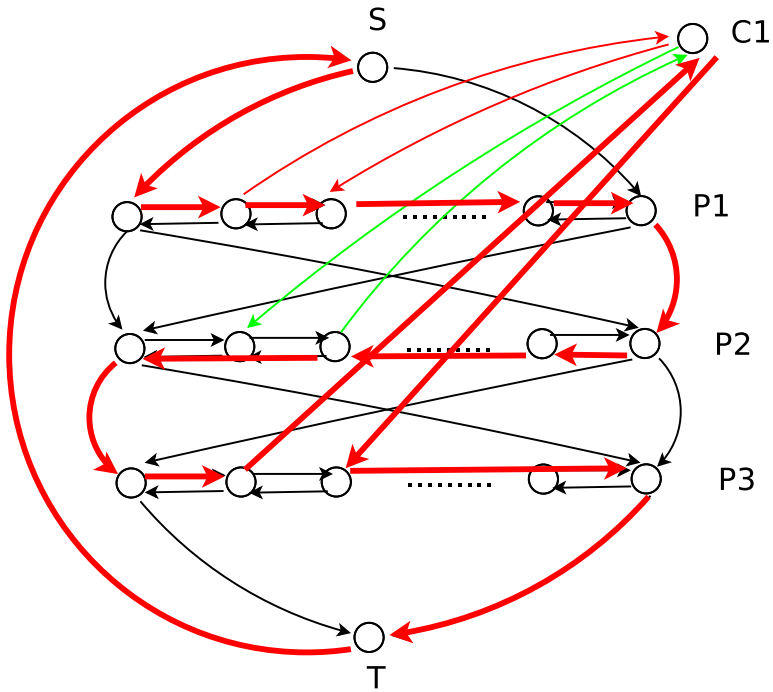


- We construct a special graph based on a SAT instance.

Does this special graph has a Hamiltonian Cycle?

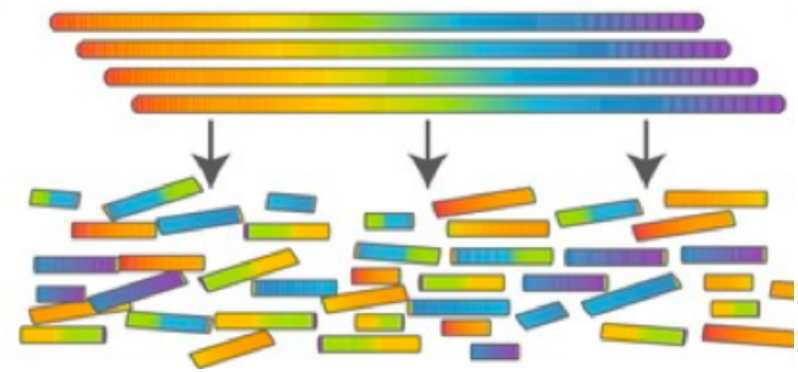
SAT Instance

(X1 OR NOT X2 OR X3)



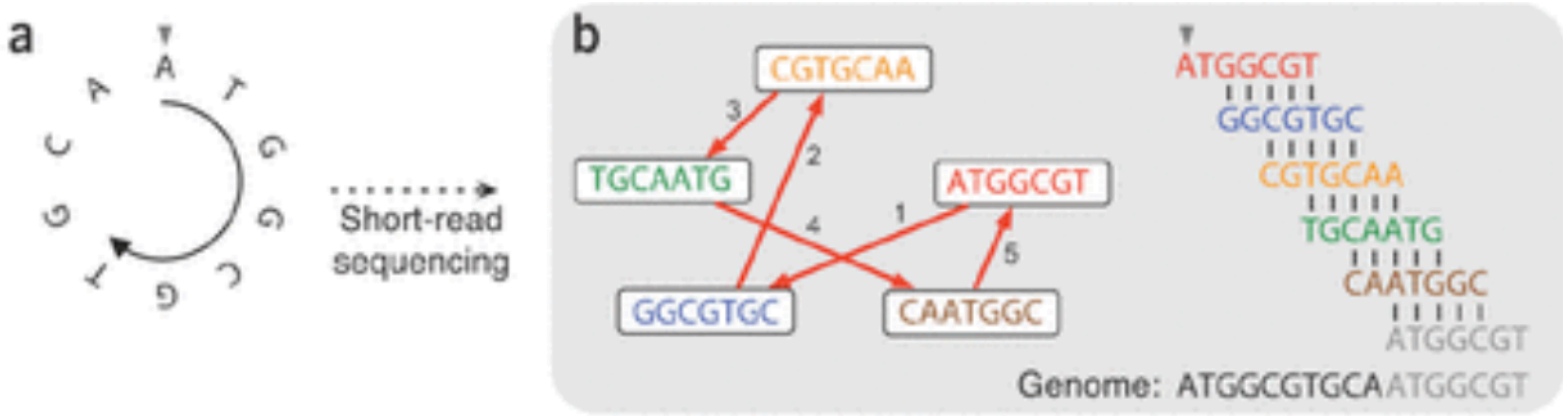
- The graph on the right panel has a Hamiltonian cycle iff the SAT instance has a true assignment.
- Thus, the Hamiltonian Cycle problem is as hard as the SAT problem, which is the hardest problem in NP.

DNA sequencing: an application of Hamiltonian Cycle



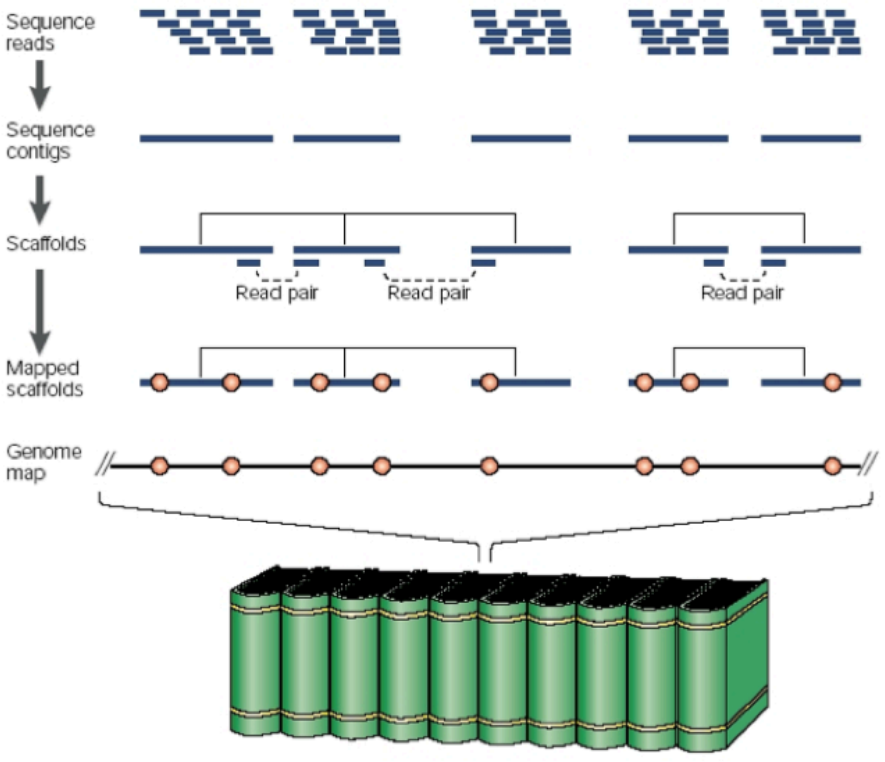
- Multiple copies of a DNA $\Rightarrow$ small sequenced fragments called reads (say 500 bp).
- Question: how to restore the whole genome from the short fragments?

Hamiltonian Cycle and genome assembly



- Let's construct a graph as follows:
  - node: a short fragment
  - edge: if two fragments overlap, then an edge is added between the corresponding nodes;
- Thus, the original genome corresponds to a Hamiltonian cycle of the graph.

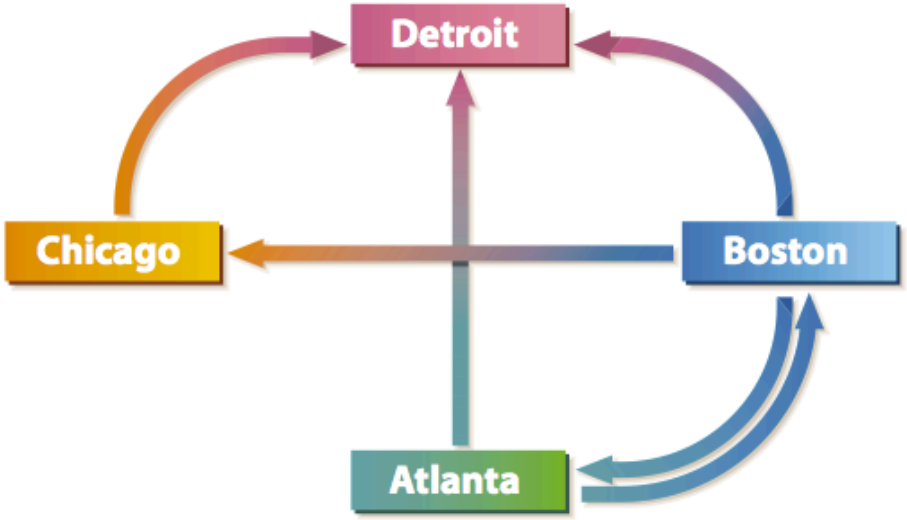
Assemble rice genome using Dawning 3000 HPC



Finding Hamiltonian Cycle using DNA computer

- In 1994, Leonard M. Adleman designed a special machine to find Hamiltonian Cycle using DNA computer.

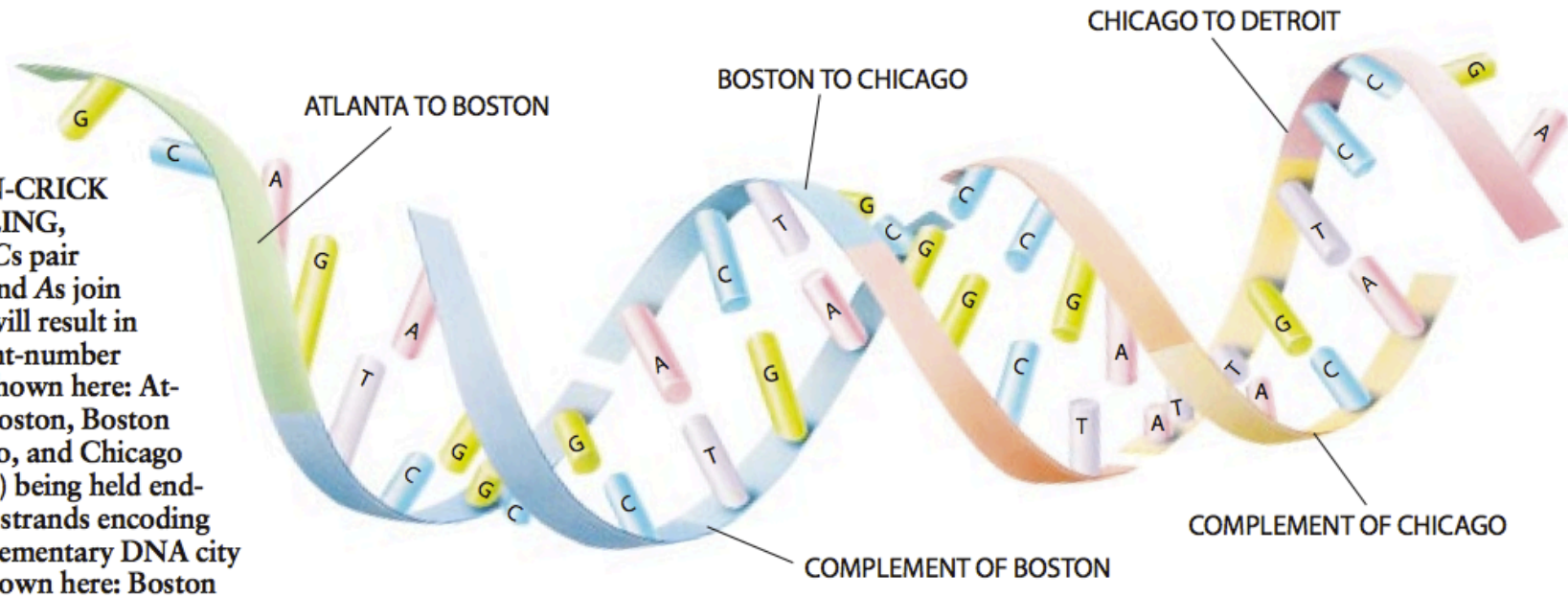
Synthezing a set of DNA fragments to represent cities and roads



CITY	DNA NAME	COMPLEMENT
ATLANTA	ACTTGCAG	TGAACGTC
BOSTON	TCGGACTG	AGCCTGAC
CHICAGO	GGCTATGT	CCGATACA
DETROIT	CCGAGCAA	GGCTCGTT

FLIGHT	DNA FLIGHT NUMBER
ATLANTA - BOSTON	GCAGTCGG
ATLANTA - DETROIT	GCAGCCGA
BOSTON - CHICAGO	ACTGGGCT
BOSTON - DETROIT	ACTGCCGA
BOSTON - ATLANTA	ACTGACTT
CHICAGO - DETROIT	ATGTCCGA

The Hamilton cycle appears when putting all DNA fragments in a tube



WATSON-CRICK ANNEALING, in which *C*s pair with *G*s and *A*s join with *T*s, will result in DNA flight-number strands (shown here: Atlanta to Boston, Boston to Chicago, and Chicago to Detroit) being held end-to-end by strands encoding the complementary DNA city names (shown here: Boston and Chicago).

Thanks