CS 3100, Models of Computation, Spring 20, L24

Ganesh Gopalakrishnan
School of Computing
University of Utah
Salt Lake City, UT 84112

cp Lec25.pptx ./Lec25

URL: bit.ly/3100s20Syllabus

Undecidability proofs

The basic idea of Mapping Reduction in Undec.

- A way to take advantage of an already proven hard-fought theorem ("A" below) in establishing the truth of new conjectures ("B" below)
  - We do a proof via diagonalization that **A_TM (H_TM) is undecidable** (the "A" mentioned)
  - We can then handle 1000s of other (often more relevant) questions easily (the "B"s)
    - Prove that "A" <= "B" i
    - i.e. A mapping-reduces-to B **via a computable function f** (i.e. you can code-up f in any of your favorite prog. languages in a way that f NEVER loops on any A-input). **This is almost always easy. Our "Translate" function was a "print" command. No need to loop in there !!**)
    - i.e. Given a **solver (or decider) for B**, we can take ANY

      A-instance, map it via "f" into a B-instance

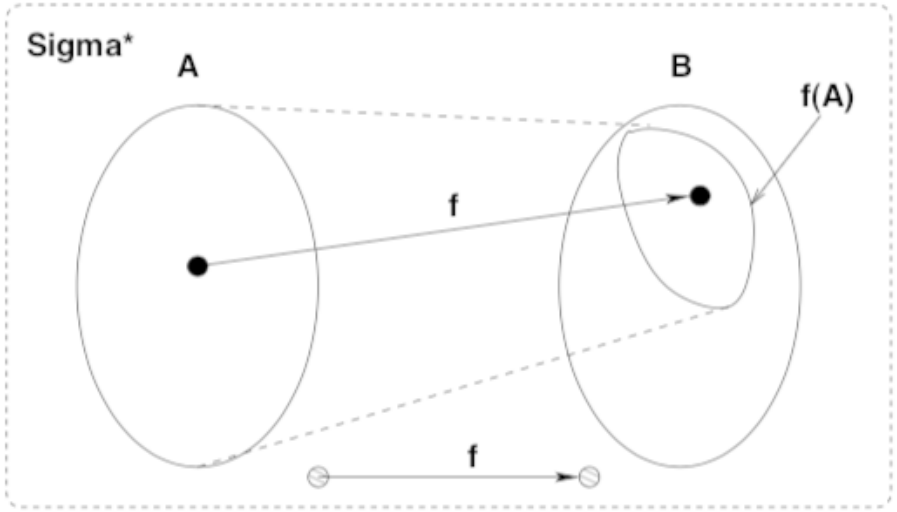
      and feed it to the claimed **B-solver (or B-decider)**

- Given that an MR means

  "x in A iff f(x) in B", i.e.
  - x in A => f(x) in B
  - x not in A => f(x) not in B

We can conclude that

"solving B means Solving A → blooper!"



NP-Completeness proofs

Important things about NPC-studies

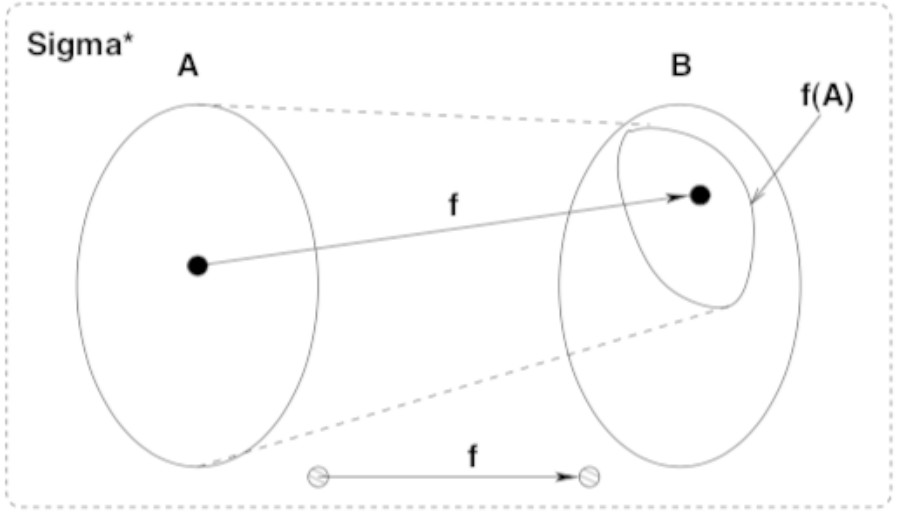
- NPC problems never loop!
  - It is always about P-time runnability
- What was “bad” (looping) is now “exp-time decider behavior (i.e. solvability in P-time is NOT KNOWN)
- NPC is a reflection of honest human ignorance
  - The only way science advances is by confessing ignorance and being brutally honest. There is ABSOLUTELY NO ROOM for saying things for the sake of situational convenience … unless facts allow things to be well-grounded
- Ever noble human endeavor must allow for retractions
  - To be perfect is to never try!

Why NPC matters (from book by Garey/Johnson)

"I can't find an efficient algorithm, but neither can all these famous people."

The basic idea of Mapping Reduction NPC proofs

- A way to take advantage of an already proven hard-fought theorem ("A" below) in establishing the truth of new conjectures ("B" below)
  - We do a proof via diagonalization that **3-SAT is NP-Complete** (the "A" mentioned)
  - We can then handle 1000s of other (often more relevant) questions easily (the "B"s)
    - Prove that "A" <= "B" i
    - i.e. A mapping-reduces-to B **via a Polynomial-time computable function f (i.e. you can code-up f in any of your favorite prog. languages in a way that f NEVER loops on any A-input and runs in P-time. Easy to do - you map Boolean formulae to Graphs, etc as we will see!)**
    - i.e. Given a **P-time solver for B**, we can take ANY A-instance, map it via "f" into a B-instance and feed it to the claimed **P-time B-solver (P-time decider)**



- Given that an MR means

"x in A iff f(x) in B", i.e.

  - x in A => f(x) in B
  - x not in A => f(x) not in B

We can conclude that

"solving B in P-time means Solving A in P-time --> A Turing Award or Clay Prize !! (not a blooper)

The language K-Clique

- Given an undirected graph, is there a set of K nodes such that they are all pairwise connected?
  - They form a K-Clique?

K-Clique – { < G > : G is a graph that has a K-Clique in it }

We are about to show 3-SAT <=p K-Clique

- <=p is a mapping reduction but with a polynomial bound on runtime
- That is, we can translate a 3-SAT instance to a K-Clique instance in polynomial time
- This means that if K-Clique has a Polynomial Algorithm, then 3-SAT will also have a Polynomial Algorithm

3SAT <=p K-Clique (the "Translate" fn.)

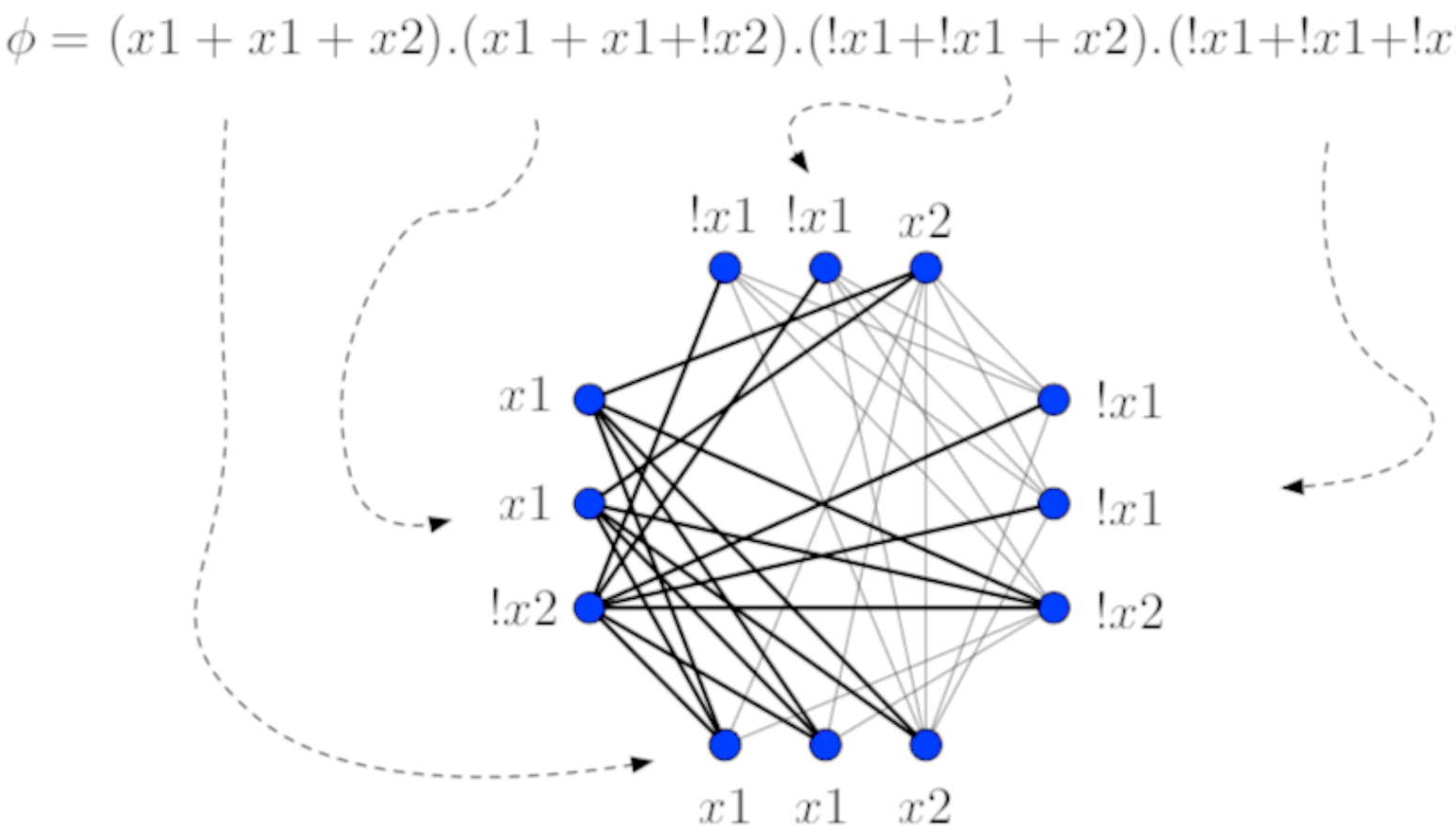


Figure 16.9: The Proof that *Clique* is NPH using an example formula $\varphi = (x1 + x1 + x2).(x1 + x1 + !x2).(!x1 + !x1 + x2).(!x1 + !x1 + !x2)$. We never connect the nodes within each clause "island" (there are four such islands, each with three nodes). Across each clause island, we draw edges in all possible ways *provided* we never connect a literal and its complement. For visual clarity, we show through dark edges all the edges emanating from the clause island for (x1+x1+!x2) going to all other clause islands. We also show the remaining edges, but using fainter lines.

Exercises to help learn 3-SAT <=p K-Clique

Practice Problems: Sat? Sat Instance? Clique?

$$\overset{=3}{(}x_1 + x_2 + x_3) \cdot (x_1 + !x_2 + !x_3)$$

Practice Problems: Sat? Sat instance? Clique?

$$\overset{=3}{(}x_1 + x_2 + x_3) \cdot (x_1 + !x_2 + !x_3)$$

Sat? Sat instance? Clique?

$$\overset{=3}{(}x_1 + x_2 + x_3) \cdot (!x_1 + !x_2 + x_3) \cdot (!x_1 + !x_2 + !x_3) \cdot (!x_1 + x_2 + !x_3)$$

Sat? Sat instance? Clique?

$$=3$$

$$(x_1 + x_2 + x_2) \cdot (!x_1 + !x_2 + !x_2) \cdot (x_1 + !x_2 + !x_2) \cdot (!x_1 + x_2 + x_2)$$

Your Asg-7's problems

- They help you practice these notions using the Binary Decision Diagram tool (BDD tool)
  - Part of Jove
- Binary Decision Diagrams will be explained now
  - They also are minimal DFA " in disguise "
- We will then study the theory of NPC armed with this knowledge

Binary Decision Diagrams

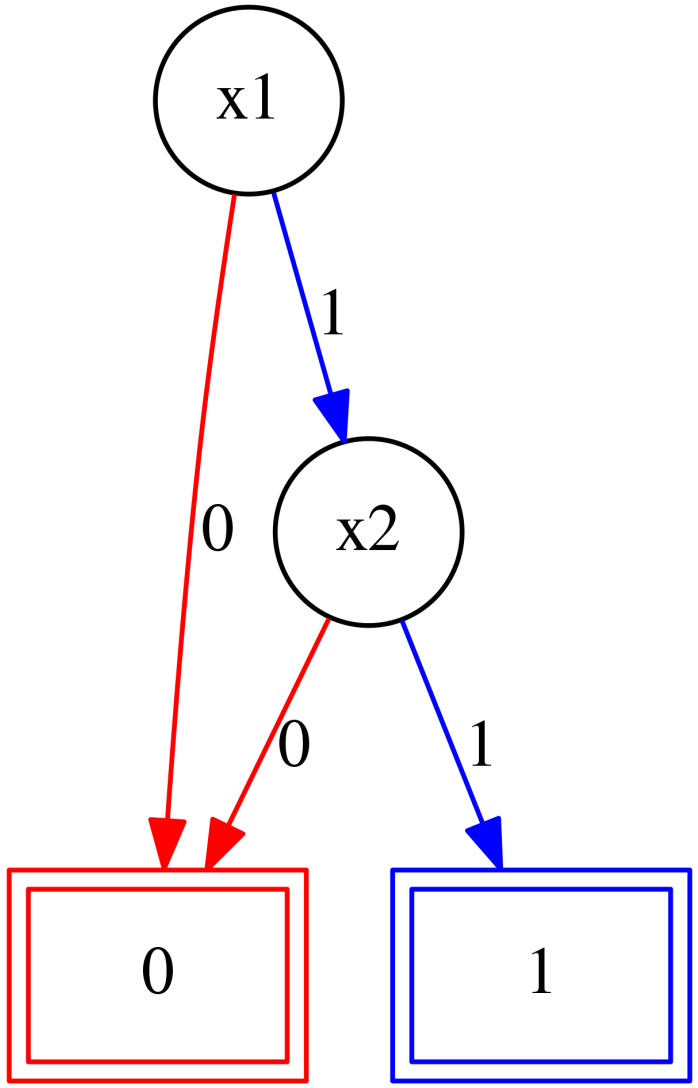
BDDs

- They are a data structure for representing Boolean Functions
- Included in Jove (see First_Jove_Tutorial/CH17/CH17.ipynb)

- All the details of BDDs is not that important
- But practice with our BDD tool will fix ideas in your mind better (seeing Boolean formulae as graphs is often edifying)

Use of BDD tool

```
# http://formal.cs.utah.edu:8080/pbl/BDD.php
# Very simple example to show-off syntax

#First declare the variables and specify variable orderin
Var_Order : x1 x2 x3

#Then define formula
fmla = (x1 | x2) & (x1 | !x2) & (!x1 | x2)

Main_Exp :  fmla

# Type “build BDD”

# Right-click and save PNG

# SAT, UNSAT, Valid – from shape of BDD
# SAT : paths exist to “1”
# UNSAT: [0] BDD
# Valid: [1] BDD
```



Things to observe about BDDs

- They are DAGs with one “layer” per variable
- They “decode” the formula in a top-to-bottom order
- The order of decoding determines the BDD size
- There are often good heuristics to select this order

Practice Problems: Sat? BDD for var order x1,x2,x3 ?

$$=^{3}(x_1 + x_2 + x_3) \cdot (x_1 + !x_2 + !x_3)$$

Sat? Sat instance with BDD var order x1,x2,x3?

$$\overset{=3}{(}x_1 + x_2 + x_3) \cdot (!x_1 + !x_2 + x_3) \cdot (!x_1 + !x_2 + !x_3) \cdot (!x_1 + x_2 + !x_3)$$

TMs help precisely model time complexity

- Here, we need to consider TMs solving problems involving Boolean expressions

Show relative hardness of problems via <=p

People observed that many real-world problems are expensive to solve in the worst case

E.g. Is there a clique in a large graph

The internet is a huge graph

Update of internet node software, reliable direct links between cities, … are all hard problems

They are all formally connected via <=p

(poly-time mapping reductions)

NPC problems are easy to check; difficult to solve

Both for 3SAT

(a + b + !b) . (!a + a + d) . (a + e + !f)

And clique

Definition of P-time and NP-time (from book)

- P-time: An algo is P-time if its computational tree is bounded in height by a polynomial function of the length of its input for every input in the language that the DTM decides. For this simple "101" DTM, here is that DTM's code and here is a computational tree – with paths shown. The paths are two for rejecting runs and one for an accepting run.

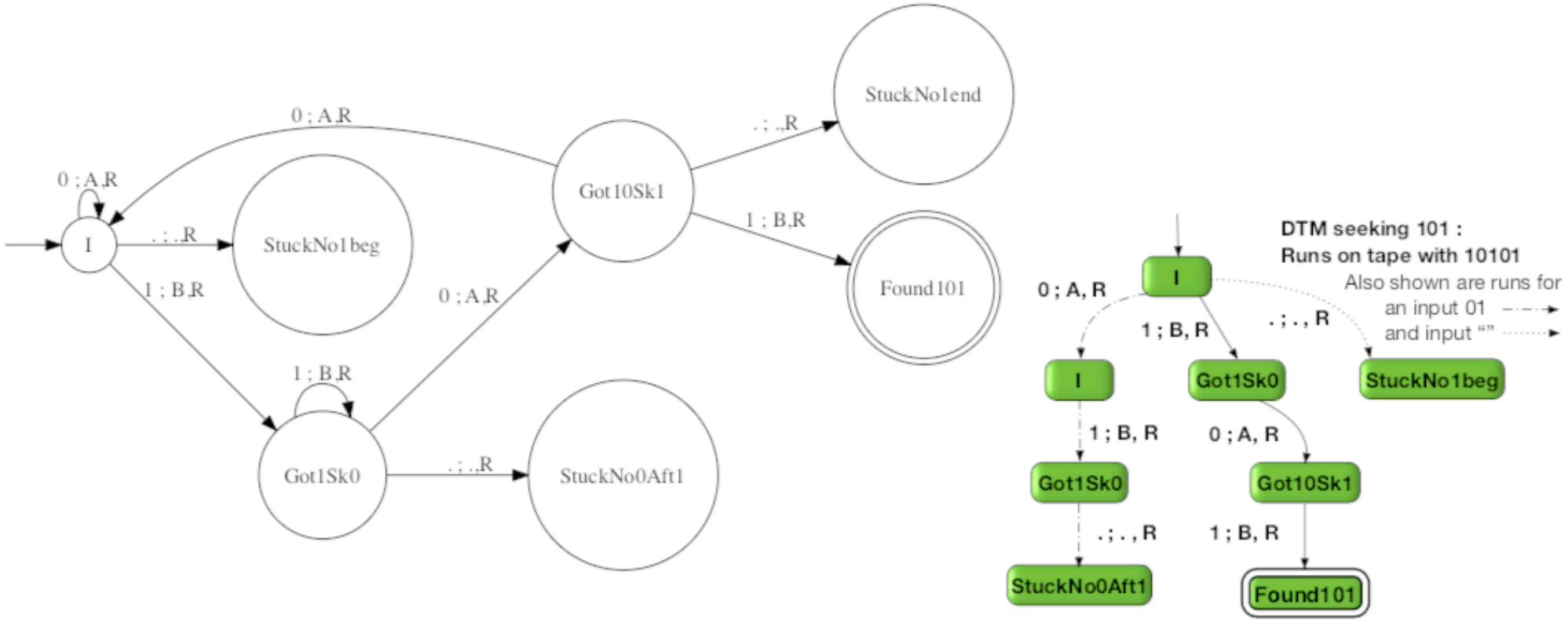
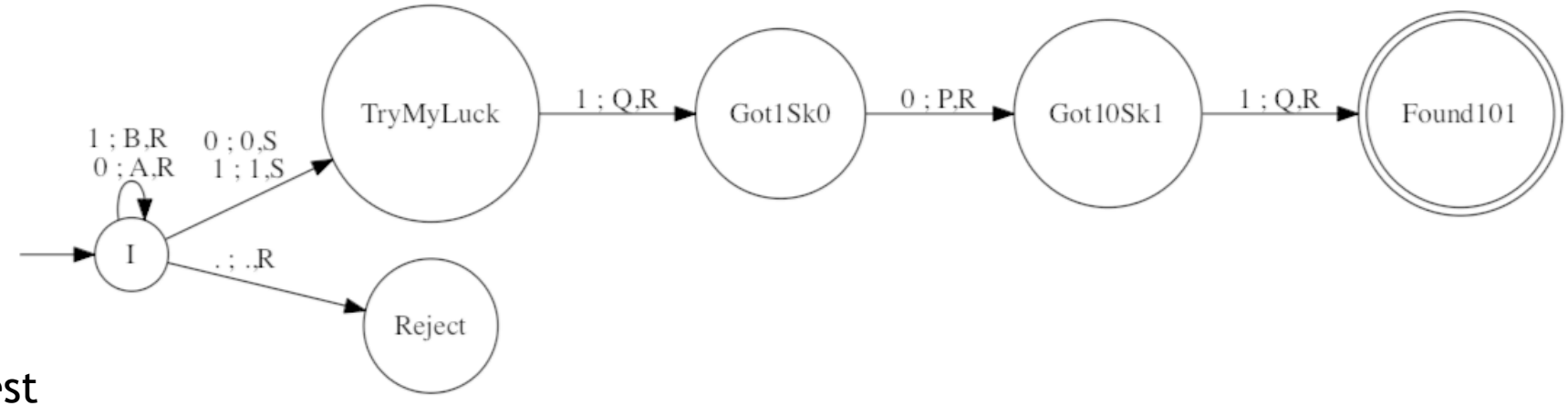


Figure 16.2: Transition diagram and computation tree for a DTM that looks for 101 within given w.

Definition of NP-time (from book)

- NP-time: An algo is NP-time if an NDTM can be obtained where it can guess a solution nondeterministically but be able to CHECK that solution in P-time. So the depth of the worst-case (deepest) path in that NDTM must be P-bounded for any input. Some problems may not even qualify for the “check phase” being P-bounded… but many useful problems have !! That makes the theory of NPC interesting and relevant in practice!

Illustration of NP-time (from book)



Look for the deepest path which accepts. That corresponds to The NP-time.

In Jove, the Fuel Models this depth Faithfully even for NDTMs…

(modulo bug-fixes if any…)

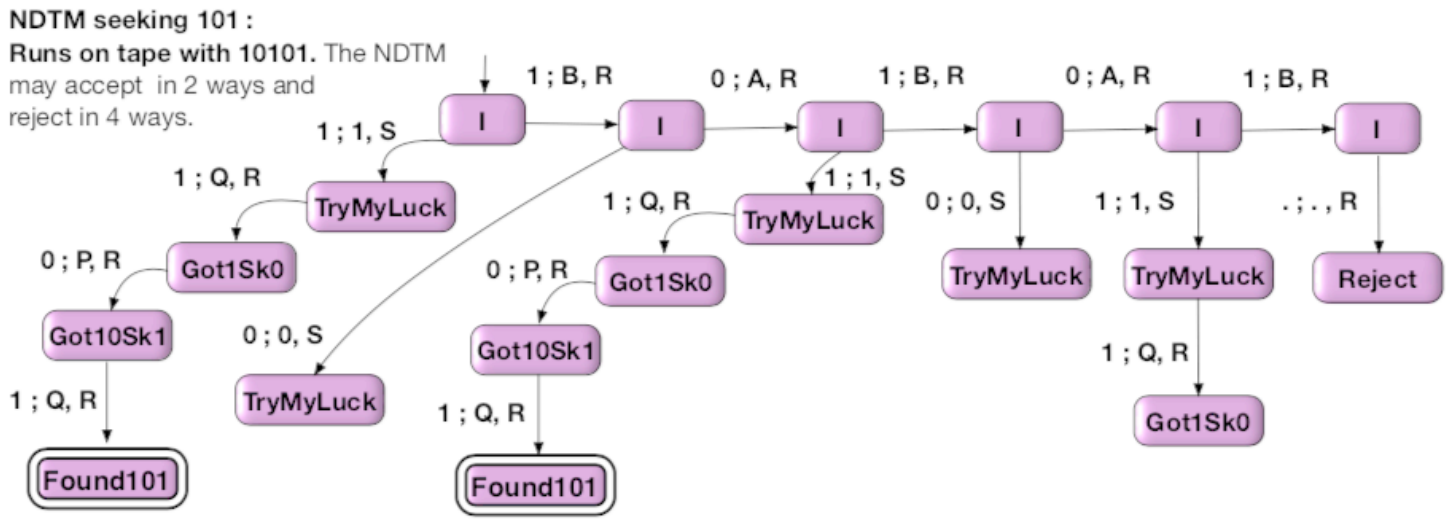


Figure 16.3: Transition diagram and computation tree for an NDTM that looks for 101 within given w.

Definition of P-time and NP-time (from book)

- P-time: Illustrated wrt 3SAT
  - Obtain computation tree of DTM for the language 3SAT (all members in it)
  - Can we claim anything about the depth of the computational tree for all inputs?
    - As far as we know, any DTM working on 3SAT appears to incur an exp depth for at least some of the instances
- NP-time: there is an NDTM that can guess the solution for a 3SAT problem (in P-time) – and also check this guess in P-time
  - Question: will we ever get a DTM that does this in P-time??
  - This is what the question of P =?= NP really means

Smart idea

- Define the idea of the hardest problems in NP
- Call it NPC

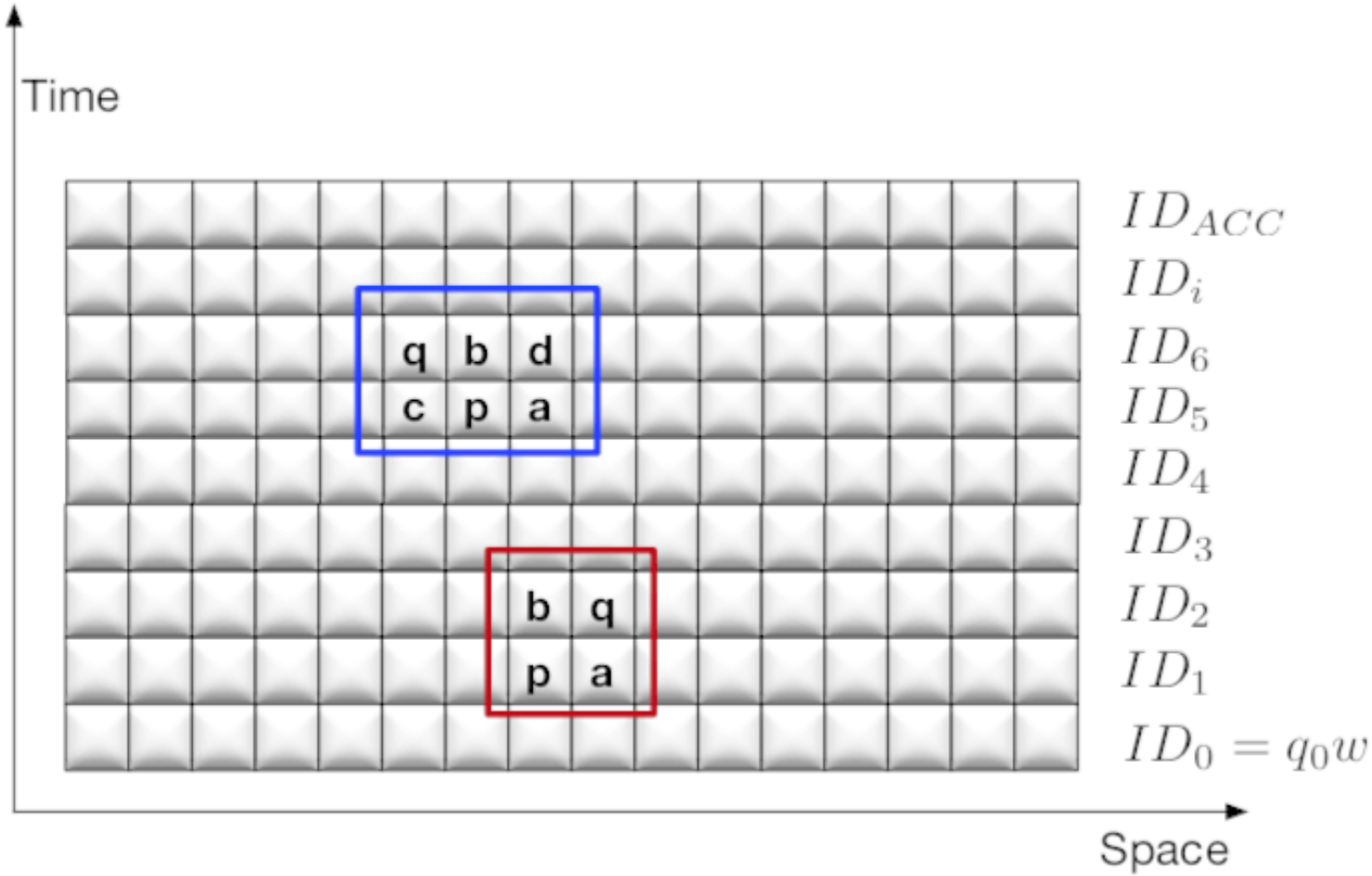
- I.e. A language L is NPC if
  - L is in NP --- has a NP-time algo (guess check on NDTP is P-time)
  - For every problem L' in NP, we have L' <=p L
    - That is, L is harder than anything there is in NP
  - Any problem such as L is "NP-hard" i.e. harder than anything in NP
- So, NPC = NP-hard + in NP
- Finding such an NPC language was the open question that Cook and Levin solved

First NPC problem

- 3SAT was shown NPC as follows
  - First show that 3SAT has an NP algorithm (easy; build an NDTM)
  - Then show that EVERY NP problem has a P-time M.R. to 3-SAT
  - This was achieved by imagining the solution of any NP problem as a collection of “tape histories”
  - Then encoding these histories using 3CNF formulae!

How 3SAT was shown NP-hard (from book)

This is the history of how a TM chugs along. Fortunately, each move from ID_k (full tape) to ID_k+1 (full tape) can be modeled as changes that occur within a 2x2 or 2x3 window. These changes can be captured using a 3CNF formula. …. The rest is history! (see the book for more)



Mapping reductions are key to “connect-up”

- NPC
  - A language L is NPC
    - If L is in NP
      - It has a P-time NDTM
    - EVERY language in NP has a P-time mapping-reduction to L

- In order to show that a NEW language L is NPC
  - We will end up producing a mapping reduction from one of the problems in NPC to L

  - **Then we have a mapping reduction from EVERY language in NP to L**

- **Study this “funnel diagram” (Ch-16) to be convinced**

The “funnel diagram”

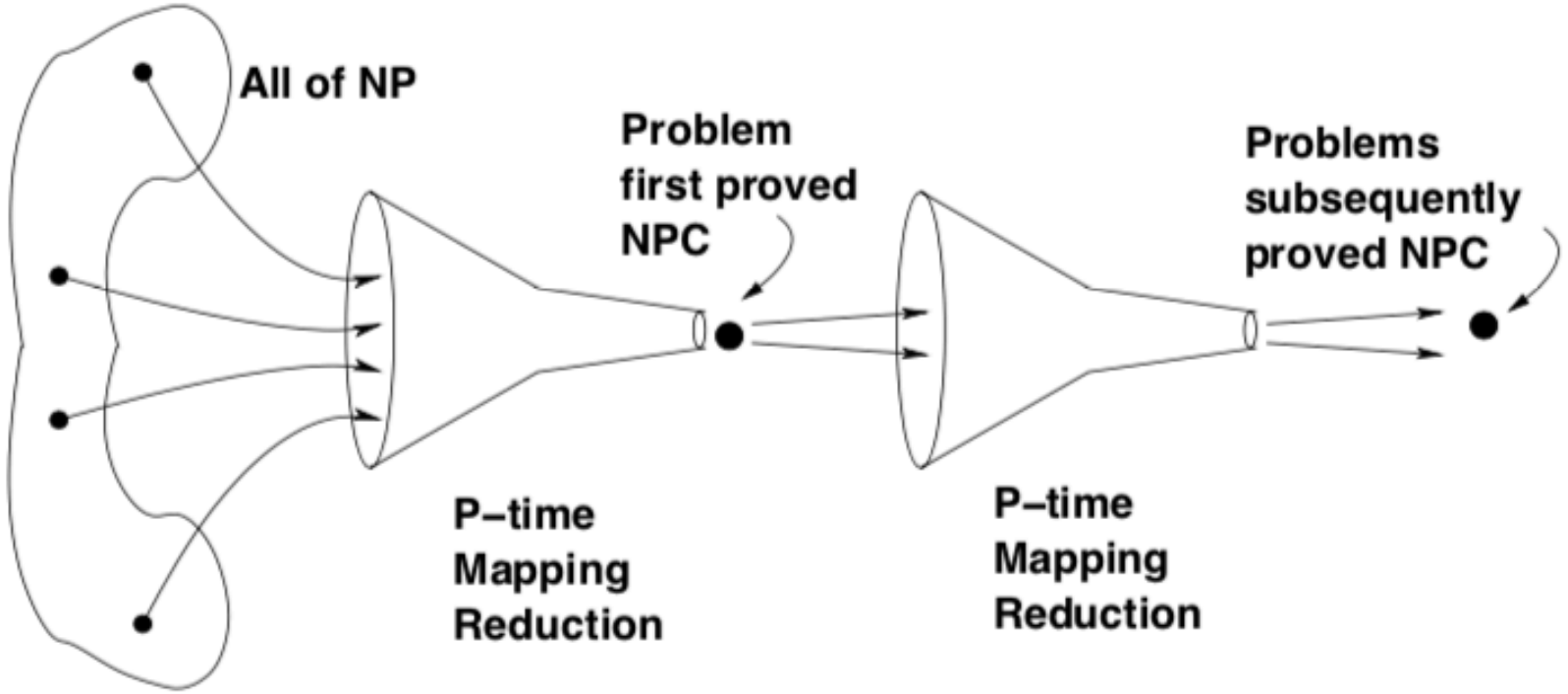


Figure 16.7: Diagram illustrating how NPC proofs are accomplished. The problem first proved NPC is 3-SAT. Definition 16.4(a) is illustrated by the “left funnel” while Definition 16.4(b) is illustrated by the “right funnel.” (The funnels serve as a gentle reminder that mapping reductions need not be onto.)

The “funnel diagram”

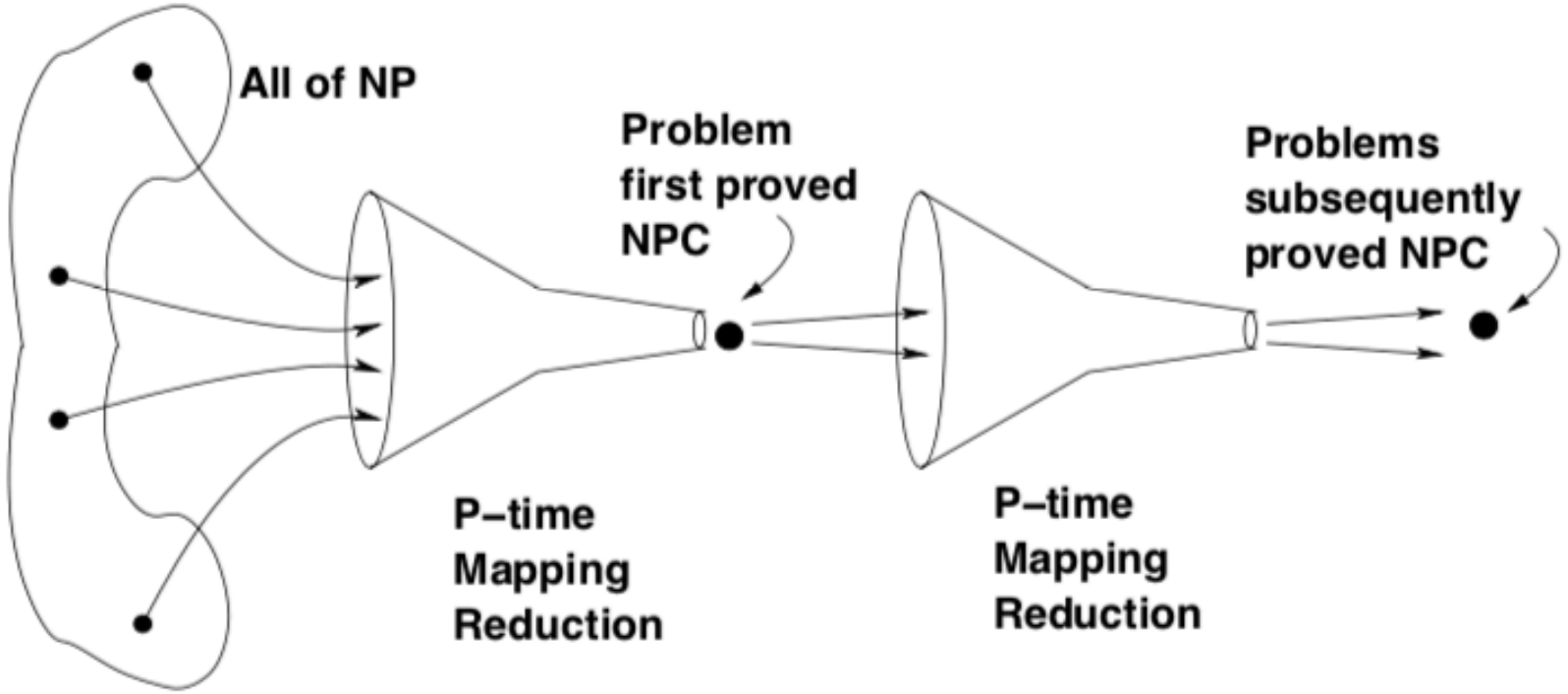
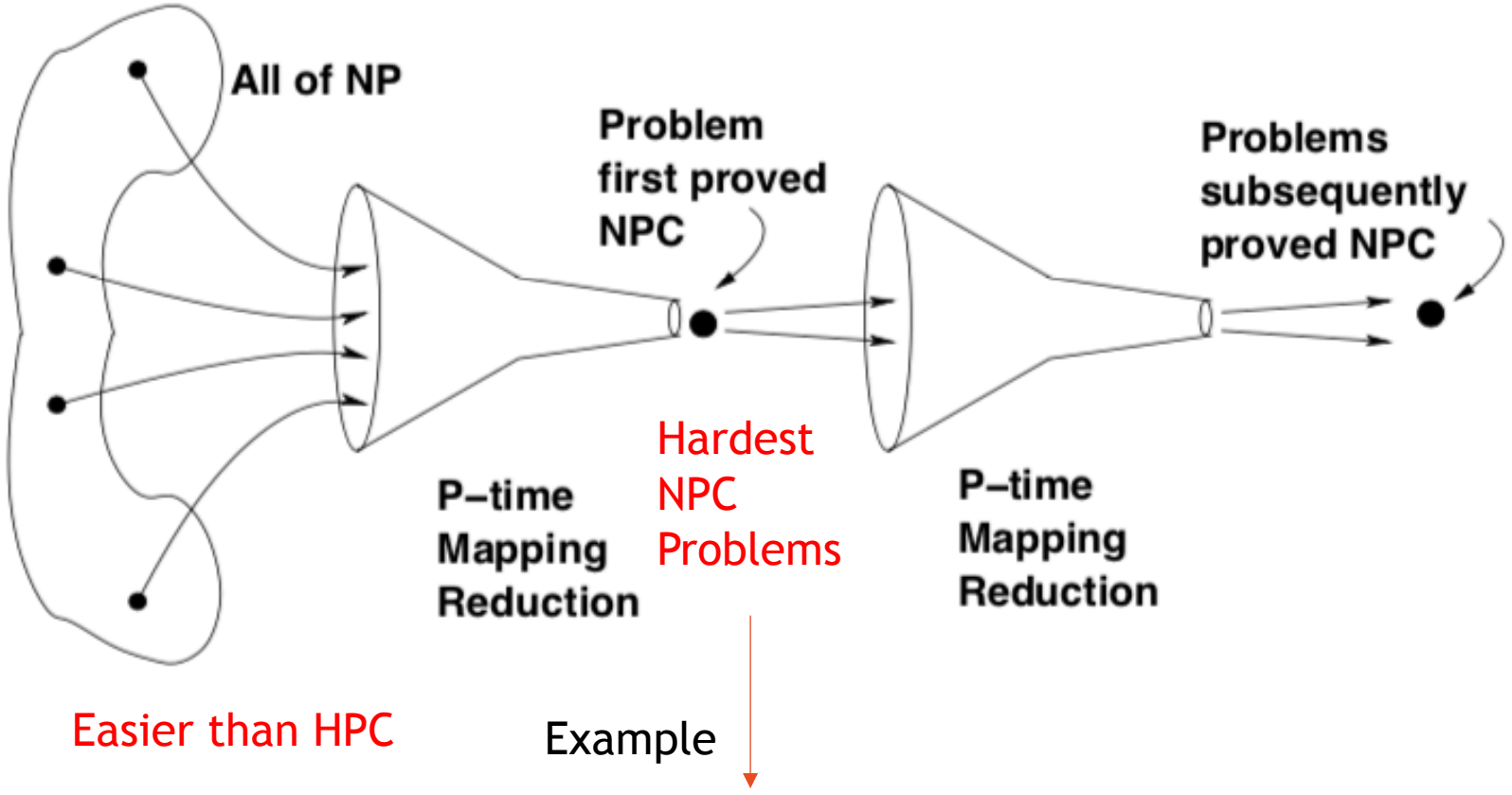


Figure 16.7: Diagram illustrating how NPC proofs are accomplished. The problem first proved NPC is 3-SAT. Definition 16.4(a) is illustrated by the “left funnel” while Definition 16.4(b) is illustrated by the “right funnel.” (The funnels serve as a gentle reminder that mapping reductions need not be onto.)

Easier than HPC

The “funnel diagram”



(a + b + !b) . (!a + a + d) . (a + e + !f)

Figure 16.7: Diagram illustrating how NPC proofs are accomplished. The problem first proved NPC is 3-SAT. Definition 16.4(a) is illustrated by the “left funnel” while Definition 16.4(b) is illustrated by the “right funnel.” (The funnels serve as a gentle reminder that mapping reductions need not be onto.)

The “funnel diagram”

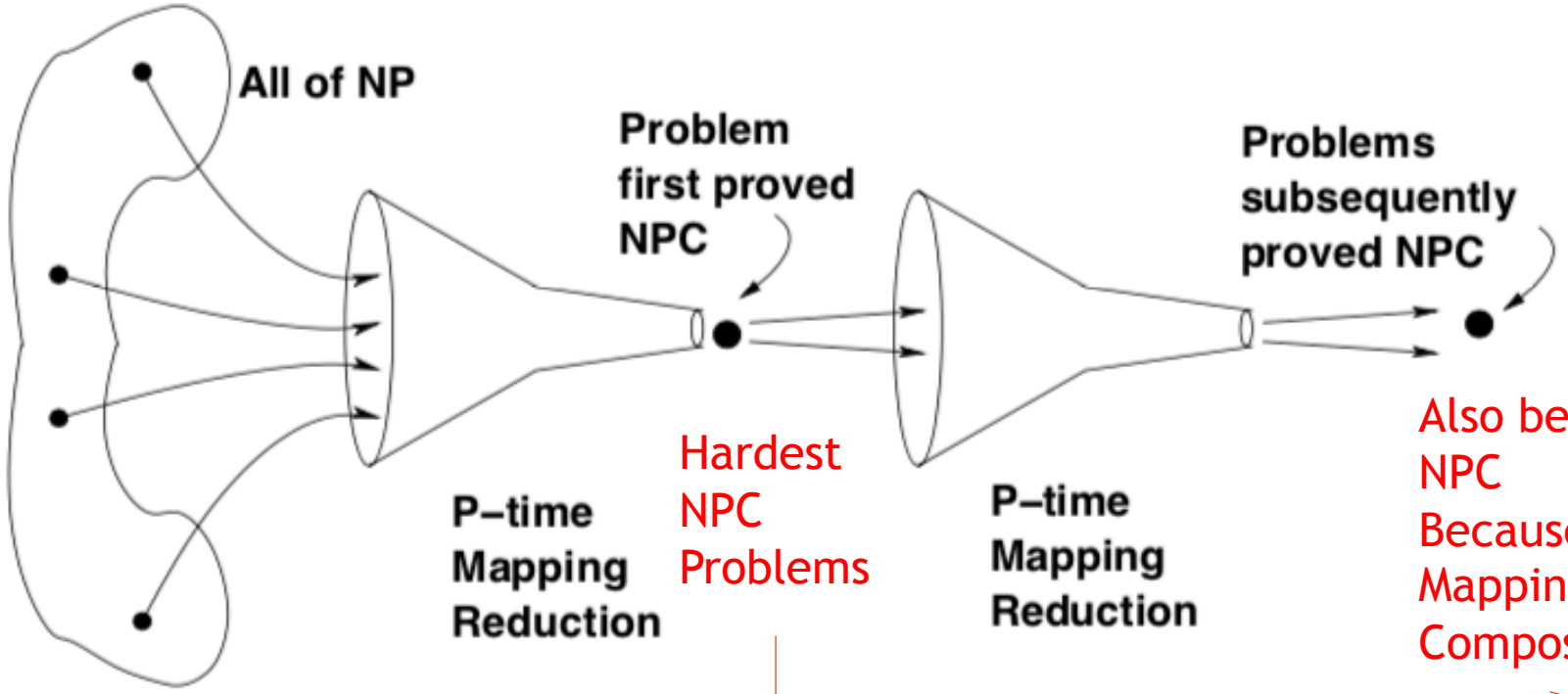


Figure 16.7: Diagram illustrating how NPC proofs are accomplished. The problem first proved NPC is 3-SAT. Definition 16.4(a) is illustrated by the “left funnel” while Definition 16.4(b) is illustrated by the “right funnel.” (The funnels serve as a gentle reminder that mapping reductions need not be onto.)

Hardest NPC Problems

Also becomes NPC Because Mapping-reductions Compose!

Easier than HPC

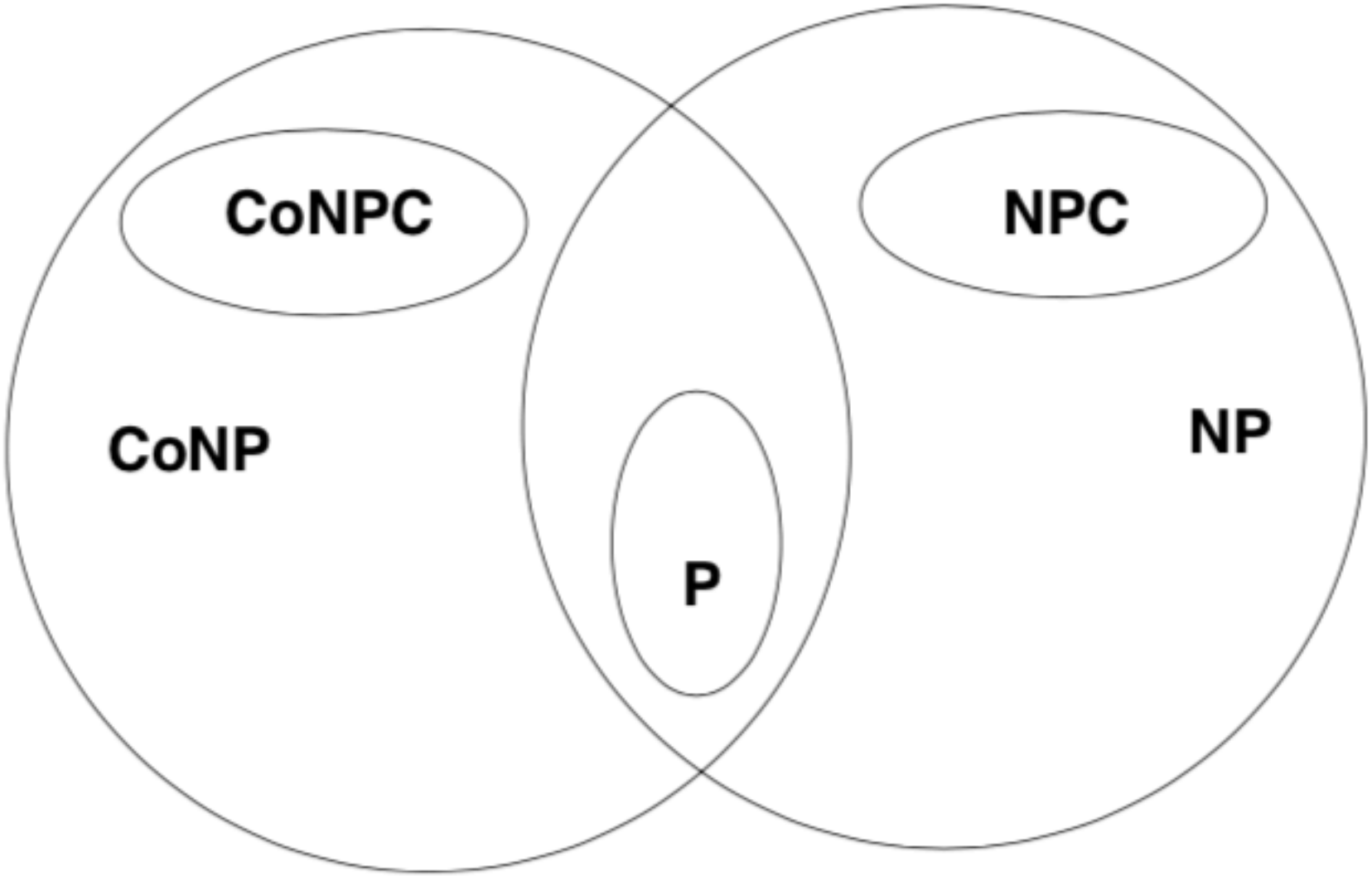
Example

(a + b + !b) . (!a + a + d) . (a + e + !f)

Example

Clique

Language hierarchy in NP-land (ignore “Co” for now)



Boolean Satisfiability: First NPC problem

- From any NDTM, we can compile a 3-SAT formula
- If the NDTM is NP-time, then we can decide the truth of the generated 3-SAT formula in NP-Time

- **If** 3-SAT is deterministic P-time, then we can decide any NDTM in P-time (not in NP-time)

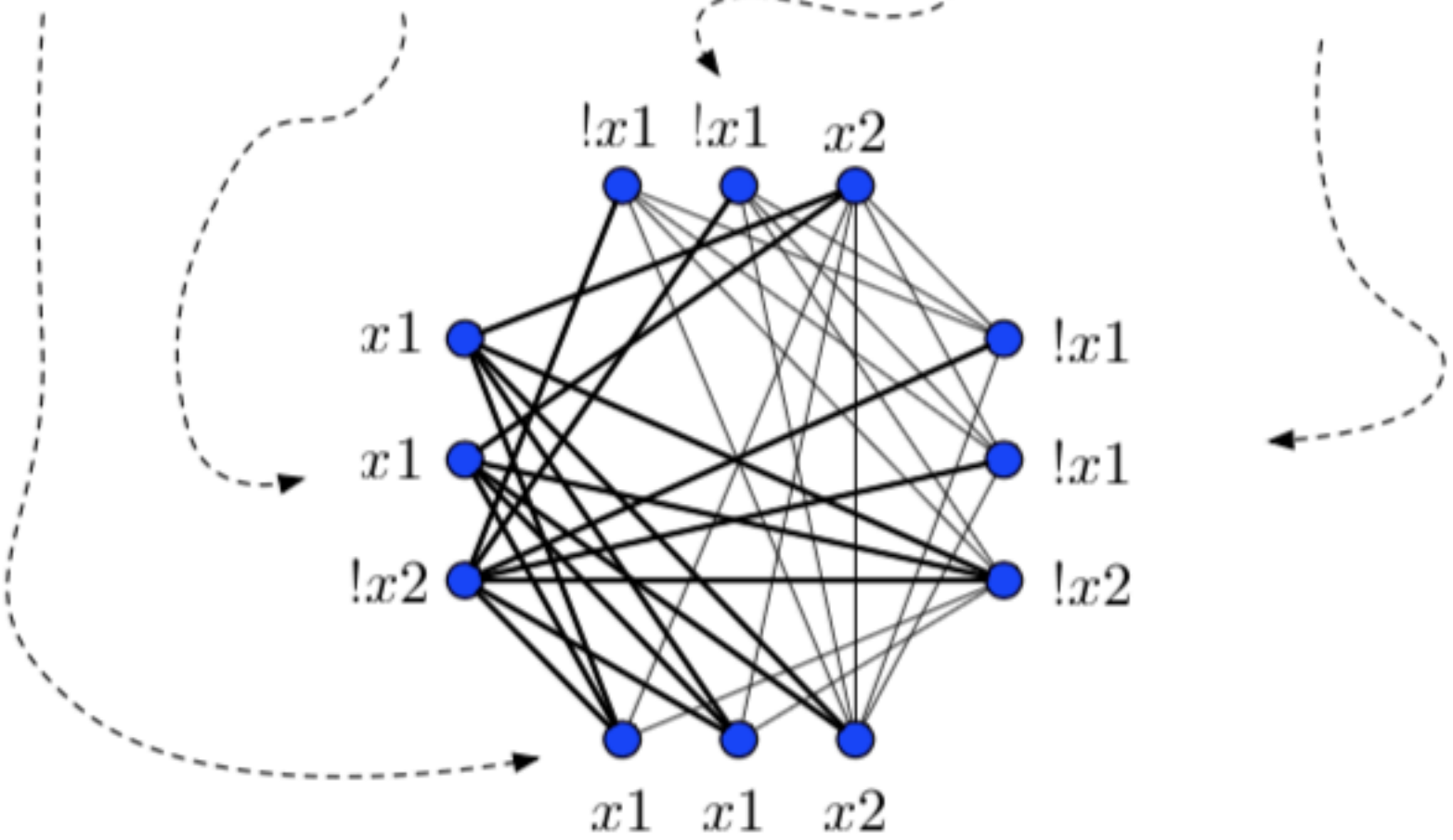
- Details given in the book

Mapping reduction based proof of K-clique being NPC

- Show K-clique is in NPC
  - **Given a description of which nodes are in the clique, we can CHECK that these nodes indeed form a K-clique**
- **Show a mapping reduction from 3-SAT of K clauses into a K-clique instance**

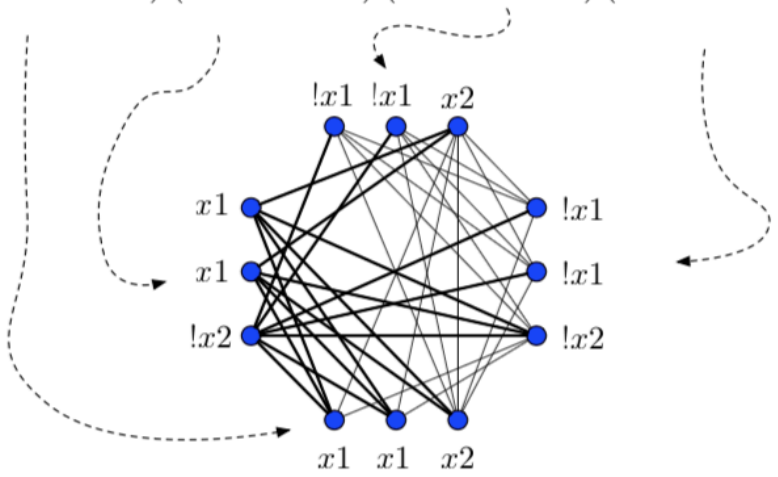
3-SAT to K-clique Mapping Reduction

$$\phi = (x1 + x1 + x2).(x1 + x1 + !x2).(!x1 + !x1 + x2).(!x1 + !x1 + !x2)$$



3-SAT to K-clique Mapping Reduction

$\phi = (x1 + x1 + x2).(x1 + x1 + !x2).(!x1 + !x1 + x2).(!x1 + !x1 + !x2)$



- Depict each clause as an “island” of 3 nodes each
- For a K-clause formula, there are K islands
- Between any two islands, draw an edge from one node of an island to another node of another island

Do the above in ALL possible ways

The edge must connect any two COMPATIBLE nodes

Nodes x and y are compatible if (x and y) is satisfiable

The given K-clause 3-CNF formula is SAT iff there is a K-clique in the mapped graph

This is a classical mapping reduction!

Thus, if there is a P-time algorithm for K-clique, then there is a P-time algorithm for 3-SAT (and ergo for all of NP)

Aside: DNF does not capture the complexity of NPC properly!

$$=3 (x_1 + x_2 + x_3) \cdot (x_1 + !x_2 + !x_3)$$

Given this or any other CNF with N variables, an NDTM can be built such that
* its first N moves are to write out a variable assignment on the tape
* then check that under that assignment, the formula is true

But hey, DNF is linear-time SAT-checkable. Multiply the above out to get a DNF

x1 + x1.!x2 + x1.!x3 + x2.x1 + x2.!x2 + x2.!x3 + x3.x1 + x3.!x2 + x3.!x3
→ then simplify

SAT if ANY product-term is ….. ?(what)

This is a linear check
But expansion to DNF turns the formula EXP LONG !!! So no real advantage.

In a sense, DNF is spelling out every possible solution and we have to check one by one !!!