RELATIONS BETWEEN COMPLEXITY CLASSES

- Basic requirements for complexity classes
- ► Complexity classes
- ► Hierarchy theorems
- ► Reachability method
- ► Class inclusions
- ► Simulating nondeterministic space
- ► Closure under complement
- (C. Papadimitriou: *Computational complexity*, Chapter 7)

C 2006 TKK, Laboratory for Theoretical Computer Science

T-79.5103 / Autumn 2006

Relations between Complexity Classes

1. Basic Requirements for Complexity Classes

A complexity class is specified by

- ► model of computation (multi-string TMs)
- mode of computation (deterministic, nondeterministic,...)
- ► resource (time, space, ...)
- \blacktriangleright bound (function f)

A complexity class is the set of all languages decided by some multi-string Turing machine M operating in the appropriate mode, and such that, for any input x, M expends at most f(|x|) units of the specified resource.

Reasonable bound functions

Definition. A function $f : \mathbf{N} \to \mathbf{N}$ is a *proper complexity function* if f is nondecreasing and there is a *k*-string TM M_f with input and output such that on any input x,

- 1. $M_f(x) = \sqcap^{f(|x|)}$ where \sqcap is a *quasi-blank* symbol,
- 2. M_f halts after O(|x| + f(|x|)) steps, and
- 3. M_f uses O(f(|x|)) space besides its input.
- **>** Examples of proper complexity functions f(n):
 - c, n, $\lceil \log n \rceil$, $\log^2 n$, $n \log n$, n^2 , $n^3 + 3n$, 2^n , \sqrt{n} , n!, ...
- ▶ If f and g are proper, so are, e.g., f + g, $f \cdot g$, 2^g .
- > Only proper complexity functions will be used as bounds.

C 2006 TKK, Laboratory for Theoretical Computer Science

T-79.5103 / Autumn 2006

Relations between Complexity Classes

4

Precise Turing machines

Definition. Let M be a deterministic/nondeterministic multi-string Turing machine (with or without input and output).

Machine *M* is *precise* if there are functions *f* and *g* such that for every $n \ge 0$, for every input *x* of length *n*, and for every computation of *M*,

- 1. *M* halts after precisely f(|x|) steps and
- 2. all of its strings (except those reserved for input and output whenever present) are at halting of length precisely g(|x|).

(Precise bounds will be convenient in various simulation results).

Simulating TMs with precise TMs	Variety of complexity classes
Proposition. Let <i>M</i> be a deterministic or nondeterministic TM	$\mathbf{P} = \mathbf{TIME}(n^k)$
deciding a language L within time/space $f(n)$ where f is proper.	
Then there is a precise TM M' which decides L in time/space $O(f(n))$.	$\mathbf{NP} = \mathbf{NTIME}(n^k)$ $\mathbf{PSPACE} = \mathbf{SPACE}(n^k)$
Proof sketch.	NPSPACE = NSPACE (n^k)
The simulating machine M'	$\mathbf{EXP} = \mathbf{TIME}(2^{n^k})$
1. computes a yardstick/alarm clock $\sqcap^{f(x)}$ using M_f and	$\mathbf{L} = \mathbf{SPACE}(\log(n))$
2. simulates M for exactly $f(x)$ steps or	$\mathbf{NL} = \mathbf{NSPACE}(\log(n))$
simulates M using exactly $f(x)$ units of space.	The relationships of these classes will be studied in the sequel.
© 2006 TKK, Laboratory for Theoretical Computer Science	© 2006 TKK, Laboratory for Theoretical Computer Science
© 2006 TKK, Laboratory for Theoretical Computer Science T-79.5103 / Autumn 2006 Relations between Complexity Classes 6	© 2006 TKK, Laboratory for Theoretical Computer Science T-79.5103 / Autumn 2006 Relations between Complexity Classes
T-79.5103 / Autumn 2006 Relations between Complexity Classes 6	T-79.5103 / Autumn 2006 Relations between Complexity Classes Complements of decision problems
T-79.5103 / Autumn 2006 Relations between Complexity Classes 6 2. Complexity Classes	T-79.5103 / Autumn 2006 Relations between Complexity Classes
T-79.5103 / Autumn 2006 Relations between Complexity Classes 6 2. Complexity Classes ► Given a proper complexity function <i>f</i> , we obtain following classes:	T-79.5103 / Autumn 2006 Relations between Complexity Classes Complements of decision problems Solution Given an alphabet Σ and a language $L \subseteq \Sigma^*$, the <i>complement</i> of
T-79.5103 / Autumn 2006 Relations between Complexity Classes 6 2. Complexity Classes Second proper complexity function f, we obtain following classes: TIME(f) (deterministic time) NTIME(f) (nondeterministic time) SPACE(f) (deterministic space)	T-79.5103 / Autumn 2006 Relations between Complexity Classes Complements of decision problems Solution Given an alphabet Σ and a language $L \subseteq \Sigma^*$, the <i>complement</i> of $\overline{L} = \Sigma^* - L$.
T-79.5103 / Autumn 2006 Relations between Complexity Classes 6 2. Complexity Classes ► Given a proper complexity function <i>f</i> , we obtain following classes: TIME(<i>f</i>) (deterministic time) NTIME(<i>f</i>) (nondeterministic time)	 T-79.5103 / Autumn 2006 Relations between Complexity Classes Complements of decision problems Given an alphabet Σ and a language L ⊆ Σ*, the <i>complement</i> of <i>L̄</i> = Σ* − L. For a decision problem A, the answer for the complement
T-79.5103 / Autumn 2006 Relations between Complexity Classes 6 2. Complexity Classes Second proper complexity function f, we obtain following classes: TIME(f) (deterministic time) NTIME(f) (nondeterministic time) SPACE(f) (deterministic space)	 T-79.5103 / Autumn 2006 Relations between Complexity Classes Complements of decision problems Given an alphabet Σ and a language L ⊆ Σ*, the <i>complement</i> of <i>L̄</i> = Σ* − L. For a decision problem A, the answer for the complement
 T-79.5103 / Autumn 2006 Relations between Complexity Classes 6 2. Complexity Classes Siven a proper complexity function f, we obtain following classes: TIME(f) (deterministic time) NTIME(f) (nondeterministic time) SPACE(f) (deterministic space) NSPACE(f) (nondeterministic space) The bound f can be a family of functions parameterized by a 	 T-79.5103 / Autumn 2006 Relations between Complexity Classes Complements of decision problems Given an alphabet Σ and a language L ⊆ Σ*, the <i>complement</i> of

Upper bound for H_f

Lemma. $H_f \in \mathbf{TIME}((f(n))^3)$.

Proof sketch.

- A 4-string machine U_f deciding H_f in time $f(n)^3$ is based on
- (i) the universal Turing machine U,
- (ii) the single-string simulator of a multi-string machine,
- (iii) the linear speedup machine, and
- (iv) the machine M_f computing the yardstick of length f(n) where n is the length of the input x.

C 2006 TKK, Laboratory for Theoretical Computer Science



Relations between Complexity Classes

12

Proof-cont'd.

The machine U_f operates as follows:

- 1. M_f computes the alarm clock $\sqcap^{f(|x|)}$ for M (string 4).
- The description of *M* is copied on string 3 and string 2 initialized to encode the initial state *s* and string 1 the input ▷*x*.
- 3. Then U_f simulates M and advances the alarm clock. If U_f finds out that M accepts input x within f(|x|) steps, then U_f accepts, but if the alarm clock expires, then U_f rejects.

Observations:

- ➤ Since *M* is simulated using a single string, each simulation step takes O(f(n)²) time.
- ▶ The total running time is $O(f(n)^3)$ for f(|x|) steps.



► For any complexity class *C*, **co***C* denotes the class

 $\{\overline{L} \mid L \in C\}.$

 All deterministic time and space complexity classes are closed under complement. Hence, e.g., P = coP.

 $\mathsf{Proof.}\xspace$ Exchange "yes" and "no" states of the deciding machine.

- The same holds for nondeterministic *space* complexity classes (to be shown in the sequel).
- ➤ An important open question: are nondeterministic *time* complexity classes closed under complement? For instance, NP = coNP?

C 2006 TKK, Laboratory for Theoretical Computer Science

Relations between Complexity Classes

T-79.5103 / Autumn 2006

3. Hierarchy Theorems

- We derive a quantitative hierarchy result: with sufficiently greater time allocation, Turing machines are able to perform more complex computational tasks.
- ▶ For a proper complexity function $f(n) \ge n$, define
 - $H_f = \{M; x \mid M \text{ accepts input } x \text{ after at most } f(|x|) \text{ steps}\}.$
- Thus H_f is the time-bounded version of H, i.e. the language of the HALTING problem.

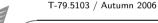
Lower bound for H_f

Lemma. $H_f \notin \mathbf{TIME}(f(\lfloor \frac{n}{2} \rfloor))$

Proof sketch.

- ► Suppose there is a TM M_{H_f} that decides H_f in time $f(\lfloor \frac{n}{2} \rfloor)$.
- ➤ Consider $D_f(M)$: if $M_{H_f}(M;M) =$ "yes" then "no" else "yes". Thus D_f on input M runs in time $f(|\frac{2|M|+1}{2}|) = f(|M|)$.
- ▶ If $D_f(D_f) =$ "yes", then $M_{H_f}(D_f, D_f) =$ "no", hence, $D_f; D_f \notin H_f$ and D_f fails to accept input D_f within $f(|D_f|)$ steps, i.e. $D_f(D_f) =$ "no", a contradiction.
- ▶ Hence, $D_f(D_f) \neq$ "yes". Then $D_f(D_f) =$ "no" and $M_{H_f}(D_f, D_f) =$ "yes". Therefore, $D_f; D_f \in H_f$, and D_f accepts input D_f within $f(|D_f|)$ steps, i.e., $D_f(D_f) =$ "yes", a contradiction again.

C 2006 TKK, Laboratory for Theoretical Computer Science



Relations between Complexity Classes

The time hierarchy theorem

Theorem. If $f(n) \ge n$ is a proper complexity function, then the class **TIME**(f(n)) is strictly contained within **TIME** $((f(2n+1))^3)$.

- ► **TIME** $(f(n)) \subseteq$ **TIME** $((f(2n+1))^3)$ as f is nondecreasing.
- ▶ By the first lemma: $H_{f(2n+1)} \in \mathbf{TIME}((f(2n+1))^3)$.
- ► By the second lemma: $H_{f(2n+1)} \notin \mathbf{TIME}(f(|\frac{2n+1}{2}|)) = \mathbf{TIME}(f(n)).$

Corollary. P is a *proper* subset of **EXP**.

- ► Since $n^k = O(2^n)$, we have $\mathbf{P} \subseteq \mathbf{TIME}(2^n) \subseteq \mathbf{EXP}$.
- ► It follows by the time hierarchy theorem that $TIME(2^n) \subset TIME((2^{2n+1})^3) \subseteq TIME(2^{n^2}) \subseteq EXP.$

The space hierarchy theorem

Theorem. If $f(n) \ge n$ is a proper complexity function, then the class **SPACE**(f(n)) is a *proper* subset of **SPACE** $(f(n)\log f(n))$.

However, counter-intuitive results are obtained if non-proper complexity functions are allowed.

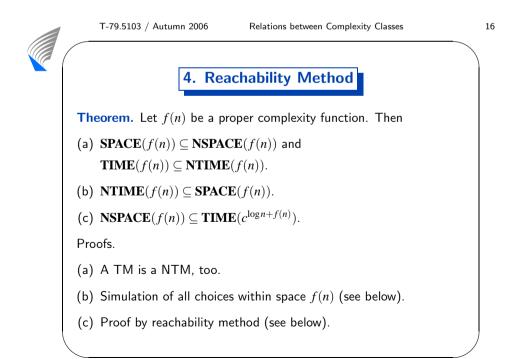
Theorem. (The Gap Theorem).

There is a recursive function f from the nonnegative integers to the nonnegative integers such that $\mathbf{TIME}(f(n)) = \mathbf{TIME}(2^{f(n)})$.

Proof sketch.

The bound f can be defined so that no TM M computing on input x with |x| = n halts after number of steps between f(n) and $2^{f(n)}$.

C 2006 TKK, Laboratory for Theoretical Computer Science



Proof of NTIME $(f(n)) \subseteq$ **SPACE**(f(n))

- ▶ Let $L \in \mathbf{NTIME}(f(n))$. Hence there is a precise nondeterministic Turing machine N that decides L in time f(n).
- Let d be the degree on nondeterminism (maximal number of possible moves for any state-symbol pair in Δ).
- ➤ Any computation of N is a f(n)-long sequence of nondeterministic choices (represented by integers 0, 1, ..., d-1).
- ➤ The simulating deterministic machine *M* considers all such sequences of choices and simulates *N* on each.

C 2006 TKK, Laboratory for Theoretical Computer Science

T-79.5103 / Autumn 2006

Relations between Complexity Classes

Proof—cont'd.

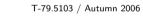
- ➤ With sequence (c₁, c₂,..., c_{f(n)}) M simulates the actions that N would have taken had N taken choice c_i at step i.
- If a sequence leads N to halting with "yes", then M does, too.
 Otherwise it considers the next sequence. If all sequences are exhausted without accepting, then M rejects.
- ➤ There is an exponential number of simulations to be tried but they can be carried out in *space* f(n) by carrying them out one-by-one, always erasing the previous simulation to reuse space.
- ➤ As f(n) is proper, the first sequence 0^{f(n)} can be generated in space f(n).

Proof of NSPACE $(f(n)) \subseteq \mathbf{TIME}(c^{\log n + f(n)})$

The *reachability method* is used to prove the claim.

- Consider a k-string *nondeterministic* TM M with input and output which decides a language L within space f(n).
- ➤ We develop a deterministic method for simulating the nondeterministic computation of *M* on input *x* within time c^{log n+f(n)} where n = |x| and c is a constant depending on *M*.
- The configuration graph G(M,x) of M is used: nodes are all possible configurations of M and there is an edge between two nodes (configurations) C₁ and C₂ iff C₁ ^M→ C₂.
- Now x ∈ L iff there is a path from C₀ = (s,▷,x,▷,ε,...,▷,ε) to some configuration of the form C = ("yes",...) in G(M,x).

C 2006 TKK, Laboratory for Theoretical Computer Science



Relations between Complexity Classes

20

Proof-cont'd.

- ➤ A configuration (q, w₁, u₁,..., w_k, u_k) is a complete "snapshot" of a computation.
- ➤ Since *M* is a machine with input and output *deciding L*:
 - the output string can be neglected,
 - for the input string, only the cursor position can change, and - for all other k-2 strings, the length is at most f(n).
- ► A configuration can be represented as $(q, i, w_2, u_2, ..., w_{k-1}, u_{k-1})$ where $1 \le i \le n$ gives the cursor position on the input string.
- ► How many possible configurations does *M* have? At most $|K|(n+1)(|\Sigma|^{f(n)})^{2(k-2)} \le |K|2n(|\Sigma|^{2(k-2)})^{f(n)} \le nc_1^{f(n)} \le c_1^{\log n+f(n)}$ for some constant *c*, depending on *M*

for some constant c_1 depending on M.

Proof—cont'd.

- ▶ Hence, deciding whether $x \in L$ holds can be done by solving a reachability problem for a graph with at most $c_1^{\log n + f(n)}$ nodes.
- ➤ The problem can be solved, say, with a quadratic algorithm in time $c_2c_1^{2(\log n+f(n))} \le c^{\log n+f(n)}$ with $c = c_2c_1^2$.
- ➤ The graph G(M,x) needs not to be represented explicitly (e.g., as an adjacency matrix) for the reachability algorithm.
- ➤ The existence of an edge from C to C' can be determined on the fly by examining C, C', and the input x.



T-79.5103 / Autumn 2006 Relations between Complexity Classes	
5. Class Inclusions	
Corollary. $L \subseteq NL \subseteq P \subseteq NP \subseteq PSPACE \subseteq EXP$.	
Proof.	
1. $\mathbf{L} = \mathbf{SPACE}(\log n) \subseteq \mathbf{NSPACE}(\log n) = \mathbf{NL}$ follows by (a).	
2. $\mathbf{NL} = \mathbf{NSPACE}(\log n) \subseteq \mathbf{TIME}(c^{\log n + \log n}) = \mathbf{TIME}(n^{2\log c}) \subseteq \mathbf{P}$ follows by (c).	
3. By (a) $\mathbf{TIME}(n^k) \subseteq \mathbf{NTIME}(n^k)$ which implies $\mathbf{P} \subseteq \mathbf{NP}$.	
4. By (b) $NTIME(n^k) \subseteq SPACE(n^k)$ which implies $NP \subseteq PSPACE$.	
5. By (a) and (c) $\mathbf{SPACE}(n^k) \subseteq \mathbf{NSPACE}(n^k) \subseteq \mathbf{TIME}(c^{\log n+n^k}) \subseteq$	
$\mathbf{TIME}(2^{n^{k+c'}}) \subseteq \mathbf{EXP}.$	

Which inclusions are proper?

Corollary. The class L is a proper subset of PSPACE.

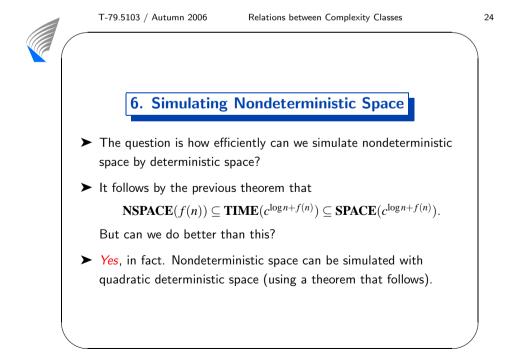
Proof. The space hierarchy theorem tells us $\mathbf{L} = \mathbf{SPACE}(\log(n)) \subset \mathbf{SPACE}(\log(n)\log(\log(n))) \subseteq \mathbf{SPACE}(n^2) \subseteq \mathbf{PSPACE}. \square$

It is believed that *all* inclusions of the complexity classes in $L \subseteq NL \subseteq P \subseteq NP \subseteq PSPACE \subseteq EXP$ are proper.

However, we only know that

- ➤ at least one of the inclusions between L and PSPACE is proper (but don't know which) and
- ➤ at least one of the inclusions between P and EXP is proper (but don't know which).

C 2006 TKK, Laboratory for Theoretical Computer Science



23

Savitch's theorem

Theorem. REACHABILITY \in **SPACE**($\log^2 n$).

Proof sketch.

- ➤ Given a graph G and nodes x, y and i ≥ 0, define PATH(x, y, i): there is a path from x to y of length at most 2ⁱ.
- If G has n nodes, any simple path is at most n long and we can solve reachability in G if we can compute whether PATH(x,y, [logn]) holds for any given nodes x, y of G.
- ► This can be done using *middle-first search*.

C 2006 TKK, Laboratory for Theoretical Computer Science

Relations between Complexity Classes

26

Proof—cont'd.

- ► function path(x, y, i) /* middle-first search */
 - if i = 0 then

T-79.5103 / Autumn 2006

if x = y or there is an edge (x, y) in G then return "yes" else for all nodes z do

if path(x,z,i-1) and path(z,y,i-1) then return "yes"; return "no"

 Proof that path(x,y,i) correctly determines PATH(x,y,i): If i = 0, then clearly path correctly determines PATH(x,y,0). For i > 0, path(x,y,i) returns "yes" iff there is a node z with path(x,z,i-1) and path(z,y,i-1) holding. By the inductive hypothesis there are paths from x to z and from z to y both at most 2ⁱ⁻¹ long. Then there is a path from x to y at most 2ⁱ long. Proof—cont'd.

- ▶ The algorithm is started with $path(x, y, \lceil \log n \rceil)$.
- ➤ O(log² n) space bound can be achieved by handling recursion using a stack containing a triple (x,y,i) for each active recursive call.
 For each node z put (x,z,i-1) into the stack and call path(x,z,i-1). If this fails, erase (x,z,i-1) and put (x,z',i-1) for the next z' otherwise erase (x,z,i-1) and put (z,y,i-1).
- ➤ As there are at most log n recursive calls active with each taking at most 3log n space, the O(log² n) space bound is achieved.

C 2006 TKK, Laboratory for Theoretical Computer Science

T-79.5103 / Autumn 2006 Relations between Complexity Classes
Corollary. For any proper complexity function f(n) ≥ log n, NSPACE(f(n)) ⊆ SPACE((f(n))²).
Proof.
To simulate an f(n)-space bounded NTM M on input x, run the previous algorithm on the configuration graph G(M,x).
The edges of the graph G(M,x) are determined on the fly by examining the input x.
The configuration graph has at most c₁^{logn+f(n)} ≤ c^{f(n)} nodes.
By Savitch's theorem, the algorithm needs at most

 $(\log c^{f(n)})^2 = f(n)^2 \log^2 c = O(f(n)^2)$ space.

Corollary. PSPACE = **NPSPACE**.

Solution Nondeterminism is less powerful with respect to space than time.

28

7. Closure under Complement

- A key result about reachability will be established: the number of nodes reachable from a node x can be computed in nondeterministic log n space!
- The complement (the number of nodes not reachable from x) can be handled in nondeterministic log n space, too!
 (This quantity can be obtained by a simple subtraction.)
- It is open (and doubtful) whether nondeterministic time complexity classes are closed under complement.

C 2006 TKK, Laboratory for Theoretical Computer Science

Relations between Complexity Classes

Functions computed by NTMs

T-79.5103 / Autumn 2006

When does a NTM M compute a function F from strings to strings?

- > On input x, each computation of M either
 - outputs the correct answer F(x) or
 - enters the rejecting "no" state.
- ➤ At least one computation must end up with *F*(*x*) which must be unique for all such computations.
- ➤ Such a machine observes a space bound f(n) iff for any input x, at halting all strings (except the ones reserved for input and output) are of length at most f(|x|).

Immerman-Szelepscényi theorem

Theorem. Given a graph G and a node x, the number of nodes reachable from x in G can be computed by a NTM within space $\log n$. Proof.

- Let us define S(k) as the set of nodes in G which are reachable from x via paths of length k or less.
- ➤ The strategy is to compute values |S(1)|, |S(2)|,..., |S(n-1)| iteratively and recursively, i.e. |S(i)| is computed from |S(i-1)|.
- Siven that the number of nodes in G is n, the number of nodes reachable from x in G is |S(n-1)|.
- ▶ Let G(v,u) mean that v = u or there is an arc from v to u in G.

C 2006 TKK, Laboratory for Theoretical Computer Science

T-79.5103 / Autumn 2006 Relations between Complexity Classes 32 Proof-cont'd. The nondeterministic algorithm: |S(0)| := 1;for k := 1, 2, ..., n - 1 do l := 0: for each node u := 1, 2, ..., n do check whether $u \in S(k)$ and set *reply* accordingly; /* See below how this is implemented */ if reply = true then l := l + 1; end for: |S(k)| := lend for

Proof–cont'd.

```
/* Check whether u \in S(k) and set reply */

m := 0; reply := false;

for each node v := 1, 2, ..., n do

/* check whether v \in S(k-1) */

w_0 := x; path := true

for p := 1, 2, ..., k-1 do

guess a node w_p; if not G(w_{p-1}, w_p) then path := false

end for

if path = true and w_{k-1} = v then

m := m+1; /* v \in S(k-1) holds */

if G(v,u) then reply := true

end if

end for

if m < |S(k-1)| then "give up" (end in "no" state)
```

© 2006 TKK, Laboratory for Theoretical Computer Science

T-79.5103 / Autumn 2006

Relations between Complexity Classes

34

Proof—cont'd.

➤ Note that only log *n*-space is needed as there are only nine variables: |S(k)|, k, l, u, m, v, p, w_p, w_{p-1}

which each (an integer) can be stored in $\log n$ space.

- The algorithm computes correctly |S(k)| (by induction on k):
 - If k = 0, then |S(k)| = 1 as given by the algorithm.
 - For k>0, consider a computation that does not "give up". We need to show that counter l is incremented iff $u\in S(k).$
 - If counter *l* is incremented, then reply = true implying that $u \in S(k)$, i.e. there is a path $(x =)w_0, \ldots, w_{k-1}(=v), u$.

If $u \in S(k)$, then there is some $v \in S(k-1)$ such that G(v,u). But as the computation does not "give up", m = |S(k-1)| (which is the correct value by induction) and therefore all $v \in S(k-1)$ are verified as such and, thus, reply is set to *true*.

– Moreover, clearly there is at least one accepting computation where paths to the members of S(k-1) are correctly guessed.

T-79.5103 / Autumn 2006

Closure under Complement

Corollary. If $f(n) \ge \log n$ is a proper complexity function, then NSPACE(f(n)) =**coNSPACE**(f(n)).

Proof sketch.

- ➤ Suppose L ∈ NSPACE(f(n)) is decided by an f(n)-space bounded NTM M. We build an f(n)-space bounded NTM M deciding L.
- ➤ On input x, M runs the previous algorithm on the configuration graph G(M, x) associated with M and x.
- ▶ \overline{M} rejects if it finds an accepting configuration in any S(k).
- ► Since G(M,x) has at most $n_g = c^{f(n)}$ nodes, then \overline{M} can accept if $|S(n_g 1)|$ is computed without an accepting configuration.

▶ Due to bound n_g , \overline{M} needs at most $\log c^{f(n)} = O(f(n))$ space.

© 2006 TKK, Laboratory for Theoretical Computer Science

