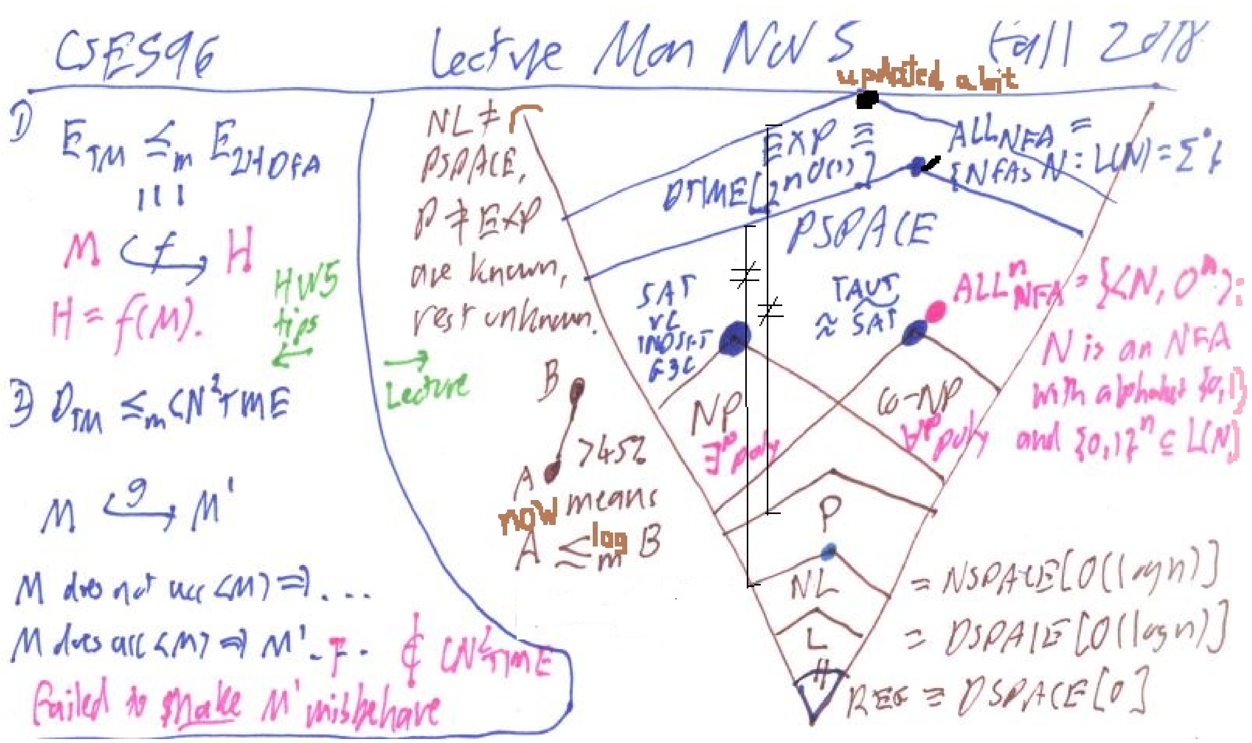
CSE491/596 Lecture Mon. Nov. 9: Time and Space Complexity Classes

From the corresponding lecture on Monday of Week 11 in 2018:



Under log-space reductions, because all languages in L are $\equiv_m^{\log}$ equivalent, the region for deterministic logspace should really warp into a point. If we could get a notion of "regular reductions" to suffice for all the NP-completeness reductions we want, we would only need to collapse REG into a point. As it stands, we can ignore little blemishes to get the big picture. It puts major classes for the four main complexity measures, viz. DTIME, DSPACE, NTIME, and NSPACE, onto one "landscape" map.

Before we go into weirdnesses like why there is no "NPSPACE" and why NL shows as closed under complements whereas NP does not, we must establish the basic "positive knowledge" about which classes are included in others. We don't have much "negative knowledge" about non-inclusions between classes of different complexity measures at all. The central mystery is why this knowledge is so much less than what we know about separations between classes defined for the *same* complexity measure. The following theorem shows the yawning exponential gaps in our current best upper bounds.

Theorem: For any "reasonable" time measure $t(n) \geq n + 1$ and space measure $s(n) \geq \log_2 n$,

$\textsf{DSPACE}[s(n)] \;\subseteq\; \textsf{NSPACE}[s(n)] \;\subseteq\; \textsf{DTIME}[2^{O(s(n)}]$

$\textsf{DTIME}[t(n)] \subseteq \textsf{NTIME}[t(n)] \subseteq \textsf{DSPACE}[O(t(n))] \subseteq \cdots$

Proof: The first and third containments are immediate by definition. For the second, let N be an NTM with some number k of tapes and work alphabet Γ that runs in space $s(n)$, and consider any input x to N, putting $n = |x|$ as usual. The notion of ``reasonable'' allows us to lay out in advance $s(n)$ tape cells that N is allowed to change. Thus any configuration I has the form $I = \langle q, w, \vec{h} \rangle$ where q is the current state, $w \in \Gamma^{s(n)}$ represents the current content of the cells N can change, and $\vec{h}$ gives the head positions on all tapes, including the location of the input head reading x. Note that I does not need to give the parts that don't change---if all cells occupied by x are kept constant, w doesn't need to include any of them. So the total number of different possible IDs we need to consider on input x is at most

$$|Q| \cdot |\Gamma|^{s(n)} \cdot (n+2)(s(n) + 2k - 2)^{k-1}.$$

Since $s(n) \geq \log_2(n)$, $|\Gamma|^{s(n)}$ is at least $2^{\log_2(n)} = n$, so the third factor does not dominate the second factor and the whole size is bounded by $2^{O(s(n))}$. (The $+2$ and $2k-2$ allow the heads to occupy blanks to the left or right of x and the cells they can change, however they are laid out on the tapes; they don't really matter to the $2^{O(s(n))}$ size estimate.)

<u>Proof of Containment (2)</u>: Let any $s(n)$ space-bounded NTM N be give

Goal: Create a $2^{O(s(n))}$-time algorithm to decide $L(N)$. Algorithm <u>M</u>:

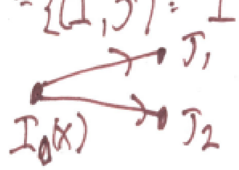
<u>↓ input x</u> <u>Wlog</u> N has a read-only input tape and s(n) cells marked off on one or more worktapes.

Build $G_x = (V, E_x)$ where

$V = \{\langle q, w, i_1, \ldots, i_k \rangle \text{ IDs}\}$

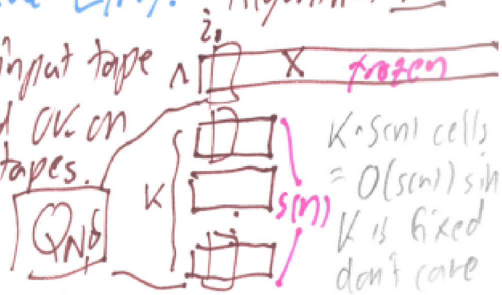
↑ contents of worktapes, not frozen x.

$E_x = \{(I, J) : I \vdash_N J\}$



N accepts x ⟺ G_x has a path from the start ID $I_0(x)$ to an accepting ID. Can tell this by running <u>BFS</u>.

How long? poly time in $\|G_x\|$ $= \|Q\| \cdot 2^{O(s(n))} \cdot n \cdot s(n)^{k-1}$ $\leq 2^{O(s(n))}$ time. ☒



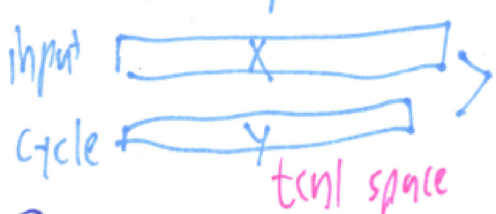
Now we define a directed graph G_x with the IDs $I, J, \ldots$ as its nodes and the relation $I \vdash_N J$ as its edge relation. Then N accepts x if and only if breadth-first search from the starting ID $I_0(x)$ finds an accepting ID (which by "good housekeeping" can be a unique node $t = I_f$). Since BFS runs in time polynomial in the size of the graph, and polynomial-in-$2^{O(s(n))}$ still gives $2^{O(s(n))}$, we obtain a deterministic algortihm that decides whether $x \in L(N)$ in time $2^{O(s(n))}$. This proves the second containment.

In-passing, we note that in the case $s(n) = O(\log n)$, DTIME$\left[2^{O(s(n))}\right]$ is just P. Also, the mapping reduction $g_N(x) = \langle G_s, I_0(x), I_f \rangle$ is computable in logspace---because we can just treat the code of any ID I as an $O(\log n)$-sized binary number and so lay out all the edges of G_x using just the δ of the fixed NTM N accepting a given language $A \in$ NL, we get that **GAP** is complete for NL under $\leq_m^{\log}$ reductions.

(1)

The fourth containment is (IMHO) best described as a *depth-first* search. Given a k-tape NTM N that runs in time $t(n)$, we may suppose N has binary nondeterminism, so that on any input x of length n there are at most $t(n)$ bits of nondeterminism that N can use. We can organize all the possible guesses y as branches of a binary tree T of depth $t(n)$ and allocate $t(n)$ cells to hold the current y we are trying. Since $N(x)$ cannot possibly use more than $kt(n)$ tape cells, we need only $t(n) + kt(n)$ space total to do a full transversal of T. We accept x if and only if an accepting branch is found. This simulation takes roughly $2^{t(n)}$ time but it all operates within $O(t(n))$ space, so $L(N) \in$ DSPACE$[O(t(n))]$.

Proof of Containment (4): Let any NTM N running in time $t(n)$ be given. Then on any input x, $|x| = n$, $N(x)$ can use at most $t(n)$ bits of nondeterminism. Hence we can cycle thru all $2^{t(n)}$ possible nondet' y's using just $t(n)$ cells to track y



Sim $N(x)$ with choices y $t(n)$ space $\rightarrow$ DTM M using $2t(n)$ space so $L(N) \in$ DSPACE$[O(t(n))]$.

$\boxtimes$

For example with $s(n) = O(\log n)$ we get

$$\mathsf{L} \subseteq \mathsf{NL} \subseteq \mathsf{P} \subseteq \mathsf{NP} \subseteq \mathsf{PSPACE}.$$

This brings us back full-circle to the deterministic space measure, and we can ratchet up to the next level:

$$\mathsf{PSPACE} \subseteq \mathsf{NPSPACE} \subseteq \mathsf{EXP} \subseteq \mathsf{NEXP} \subseteq \mathsf{EXPSPACE}.$$

A reminder again that $\mathsf{EXP} = \mathsf{DTIME}[2^{n^{O(1)}}]$, not $\mathsf{DTIME}\left[2^{O(n)}\right]$, which is called E. We do in fact have $\mathsf{PSPACE} = \mathsf{NPSPACE}$ by **Savitch's Theorem**, which we will prove next week in-tandem with showing that the language of true quantified Boolean formulas (called **TQBF** or, confusingly, **QBF**) is complete for PSPACE under $\leq_m^{\log}$ reductions. But basically the only multi-link chain of *differences* that we know from these classes is

$$\mathsf{NL} \subsetneq \mathsf{PSPACE} \subsetneq \mathsf{EXPSPACE}.$$

One can put L in place of NL here, and also prepend REG as a fourth proper link, but the main fact is the two exponential gaps thus-far seemingly needed to climb back around to the deterministic or nondeterministic space measure. Of final note today is the following theorem.

Immerman-Szelépcsenyi Theorem: For every space measure $s(n) = \Omega(\log n)$, $\mathsf{NSPACE}[s(n)]$ is closed under complements. In particular, $\mathsf{NL} = \text{co-}\mathsf{NL}$, and for linear space, $\mathsf{NLBA} = \text{co-}\mathsf{NLBA}$.

This was proved independently by Neil Immerman and Robert Szelépcsenyi in 1988. The proof is difficult and skipped here.

[The next lecture will cover the deterministic Time and Space Hierarchy Theorems in-tandem using a technical lemma from the notes https://cse.buffalo.edu/~regan/cse491596/CSE596inclusions.pdf]