

7. 時間限定チューリングマシンと クラスP

7-1. 入力サイズ

定義 : (問題の) 入力サイズ

チューリングマシンの入力記号 x の長さ $|x|$ を
入力サイズという。

通常は、入力サイズとしてはもっとも短い表現での
長さが利用される。

例えば、次のような合成数の問題における入力サイズは、
インスタンスの桁数、
すなわち、 $\log n$ を入力サイズとみなすことが多い。

問題

名称: 合成数の問題

インスタンス: 整数 n

問: n は合成数か?

問題

名称: 合成数の問題

インスタンス: 整数 n

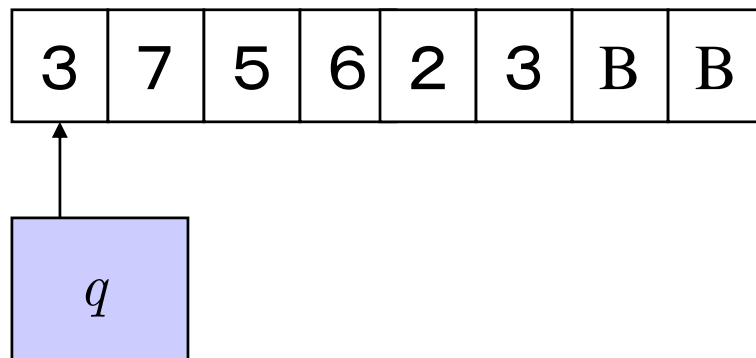
問: n は合成数か?

このインスタンスをチューリングマシンで表現することを考える。

インスタンスとして、

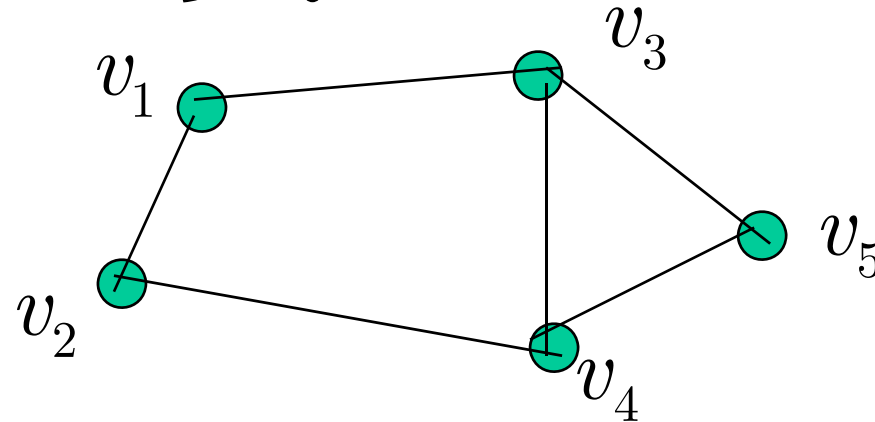
375623

のようなものを調べるときには次のような状態から行う。



よって、入力サイズは、 $|375623| = 6$ である。

一方、グラフの問題では、頂点数や、辺数を入力サイズとみなすことが多い。



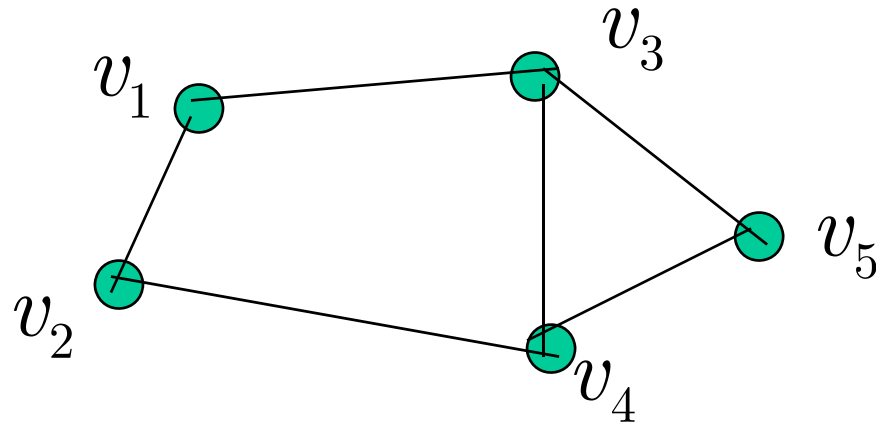
$G=(V,E)$ には、ハミルトン閉路があるのか。

問題

名称: ハミルトン閉路

インスタンス: グラフ $G = (V, E)$

問: G には、ハミルトン閉路が存在するか?



このインスタンスをチューリングマシンに入力する。

$\{(v_1, v_2), (v_1, v_3), (v_2, v_4), (v_3, v_4), (v_3, v_5), (v_4, v_5),\}$

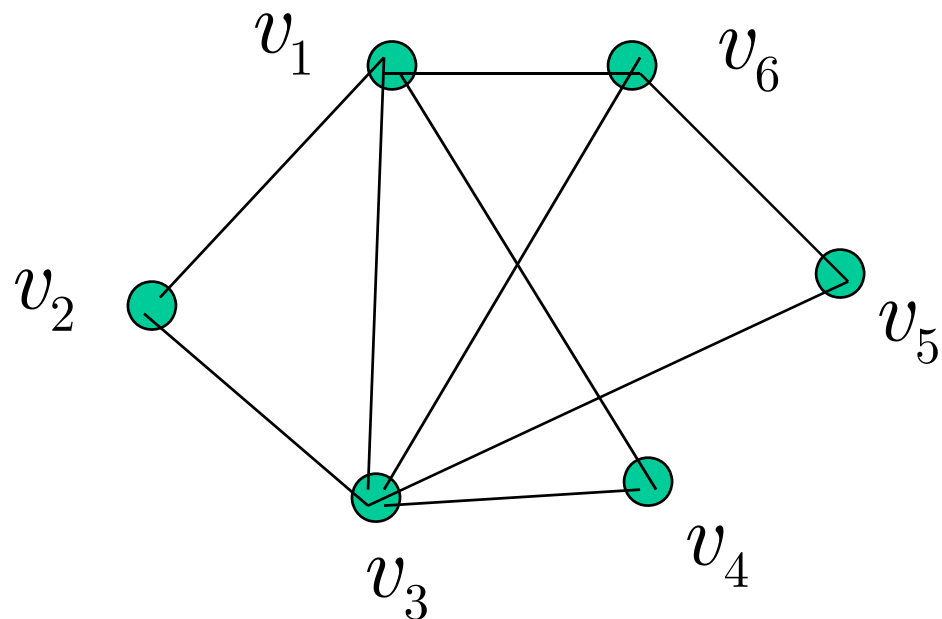
という文字列で表現可能である。よって、
この入力サイズは、

$$\left| \{(v_1, v_2), (v_1, v_3), (v_2, v_4), (v_3, v_4), (v_3, v_5), (v_4, v_5),\} \right|$$

$$\leq 6m + 2$$

$$= 6 | E | + 2$$

練習



次のインスタンスをチューリングマシンに入力形式にして、入力サイズを求めよ。ただし、入力アルファベット Γ は、以下のように与えられるものとする。

$$\Gamma = \{ "((", ")", "(", ")", "v_1, v_2, v_3, v_4, v_5, v_6" \}$$

7-2. 時間限定チューリングマシン

定義 : 時間限定チューリングマシン

入力サイズが n のとき、

$T(n)$ 時間限定チューリングマシンとは、
 $T(n)$ 回の遷移後(ステップ後)に
必ず停止する多テープチューリングマシンである。

テープ数と計算時間

次の言語を認識する2テープチューリングマシンを考える。

$$L = \{ xcx^R \mid x \in \{a, b\}^* \}$$

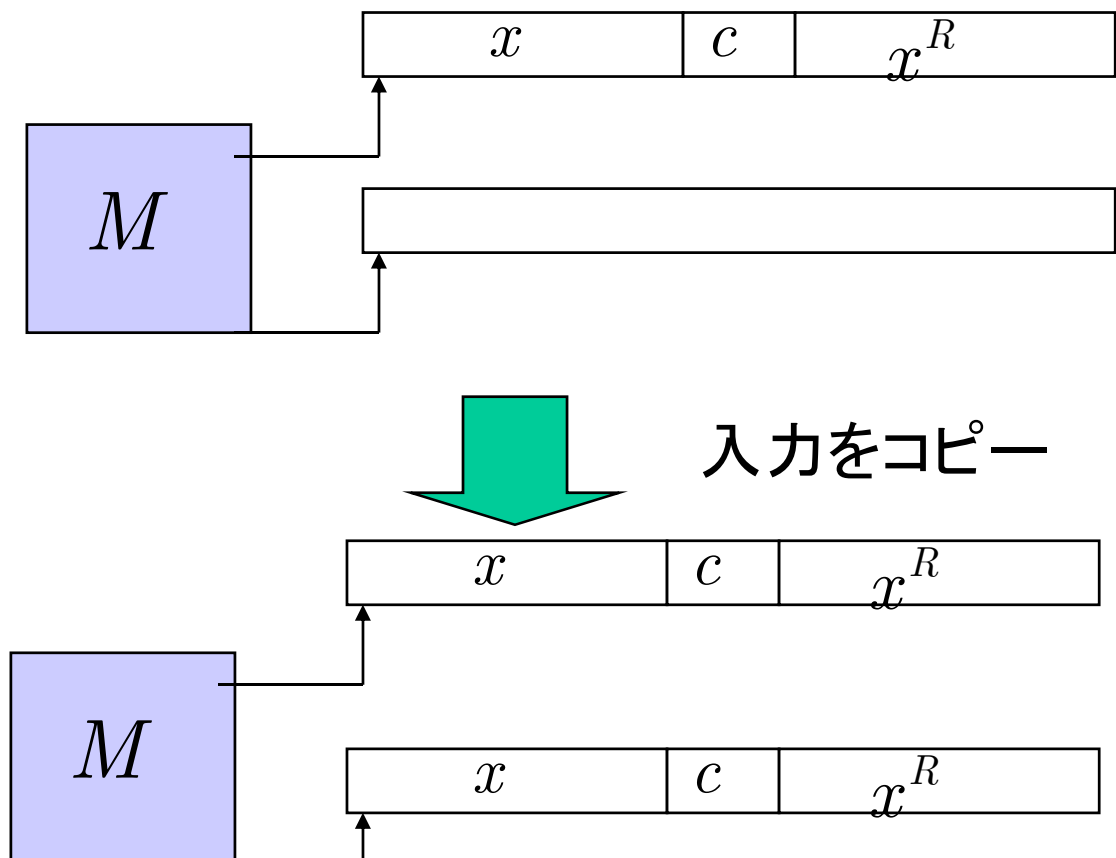
命題

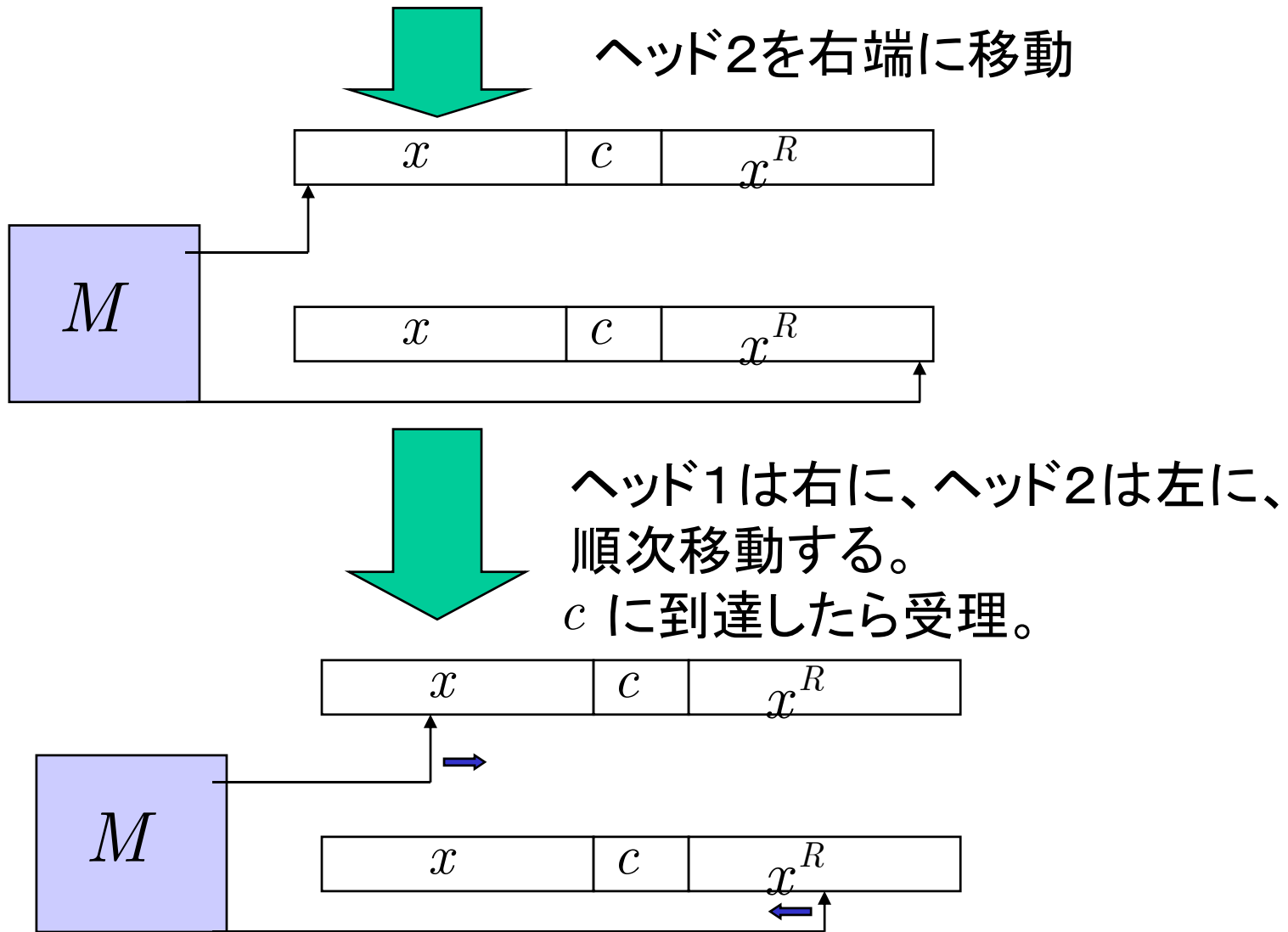
$n = |xcx^R|$ とする。このとき、

Lを認識する $O(n)$ (時間限定)2テープ
チューリングマシンが存在する。

証明

Lを認識する2テープTMを構成する。





これらはすべて ステップで行える。
したがって、 $O(n)$ 時間で認識可能。

命題

$L = \{xcx^R \mid x \in \{a,b\}^*\}$ とし、

$n = |xcx^R|$ とする。このとき、

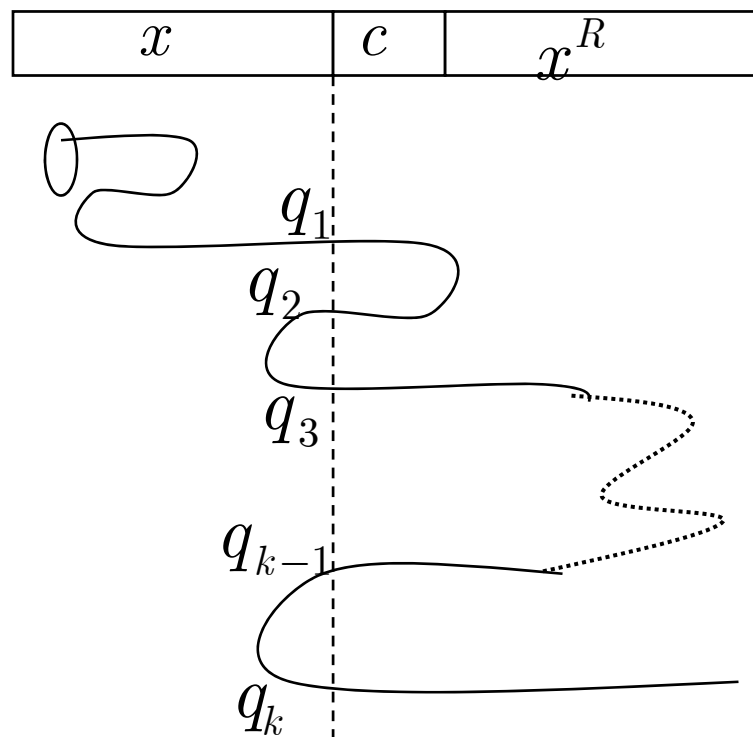
1テープTMでは、

$O(n^2)$ -TMしか存在しない。

証明

証明のために、**交差列**というアイデアを用いる。
交差列とは、一つのセルをヘッドが横切るときの
TMの状態の系列である。

交差列



ヘッド移動の様子

セル c に置ける交差列

$$q_1 q_2 \cdots q_k$$

問題の言語を認識する1テープTMをMとし、その受理動作を観察する。

テープ中央のセル c における交差列を

$$q_1 q_2 \cdots q_k$$

とする。

ここで、簡単のために、Mの状態数を8と仮定する。
(これは、有限であればいくらでもよい。議論の簡単化のために設定したにすぎない。)

このとき、交差列の種類は、

$$8^k + 8^{k-1} + \cdots + 8 \leq 8^{k+1} = 2^{3k+3}$$

である。

一方、

長さ $\frac{n}{2}$ の文字列の種類は、 $2^{\frac{n}{2}}$ 通りである。

ここで、これらの事実から、

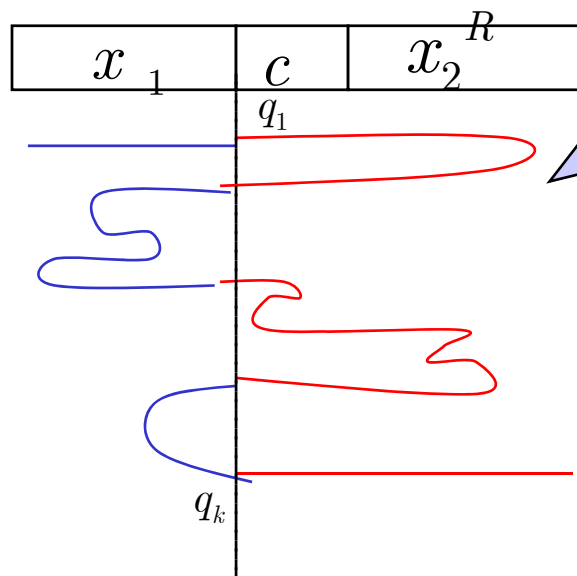
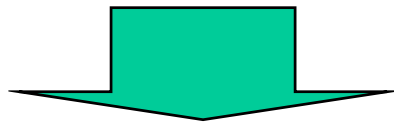
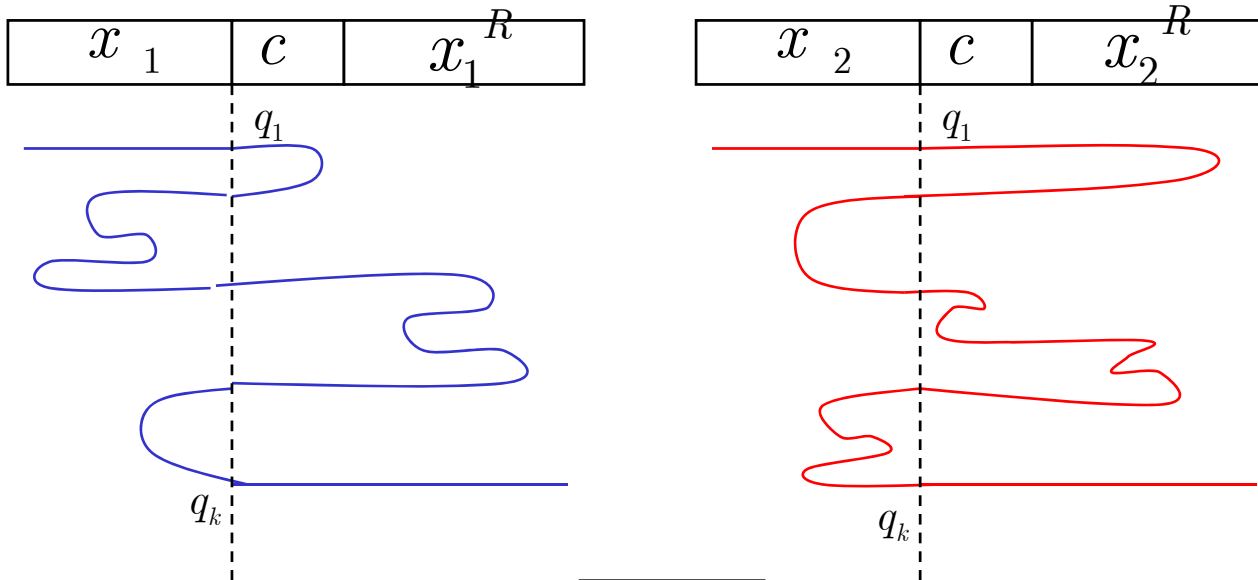
$$n \leq 6k - 6$$

であることを示す。

$$n > 6k - 6$$

とすると、文字列 xCx^R の数が
交差列 $q_1q_2 \cdots q_k$ の種類を超えてしまう。

これは、一つの交差列に対して、
2つの異なる文字列が対応することになる。
それらの文字列を $x_1Cx_1^R$ 、 $x_2Cx_2^R$ とする。



これは、
 $x_1 c x_2^R$

も認識してしまう。
 矛盾である。

したがって、テープ中央のセルにおける交差列は、

$$\frac{n}{6}$$

以上である。

テープ中央以外の交差列も類似の議論が行える。

左から、 i コマ目の交差列は、 $\frac{i}{3}$ 以上である。
よって、すべての場所の交差列の長さの和は、

$$2 \cdot \frac{1}{3} \cdot (1 + 2 + \dots + \frac{n}{2}) \geq \frac{n^2}{12}$$

以上である。

交差列の和がステップ数であるので、

結局 $O(n^2)$ 時間以上必要である。

7-3. 記号の増加による高速化技法

ここでは、時間限定チューリングマシンにおいて、定数倍は意味があまりないことを示す。

性質

$T(n) > O(n)$ とする。

このとき、ある問題が $T(n)$ 時間限定 k テープ TM で解けるなら、その言語は、
任意の定数 $c < 1$ に対して、
 $cT(n)$ 時間限定 k テープ TM で解ける。

証明

2テープTMを考える。(一般の k についても同様に行える。)

M を与えられたTMとする。

このとき、

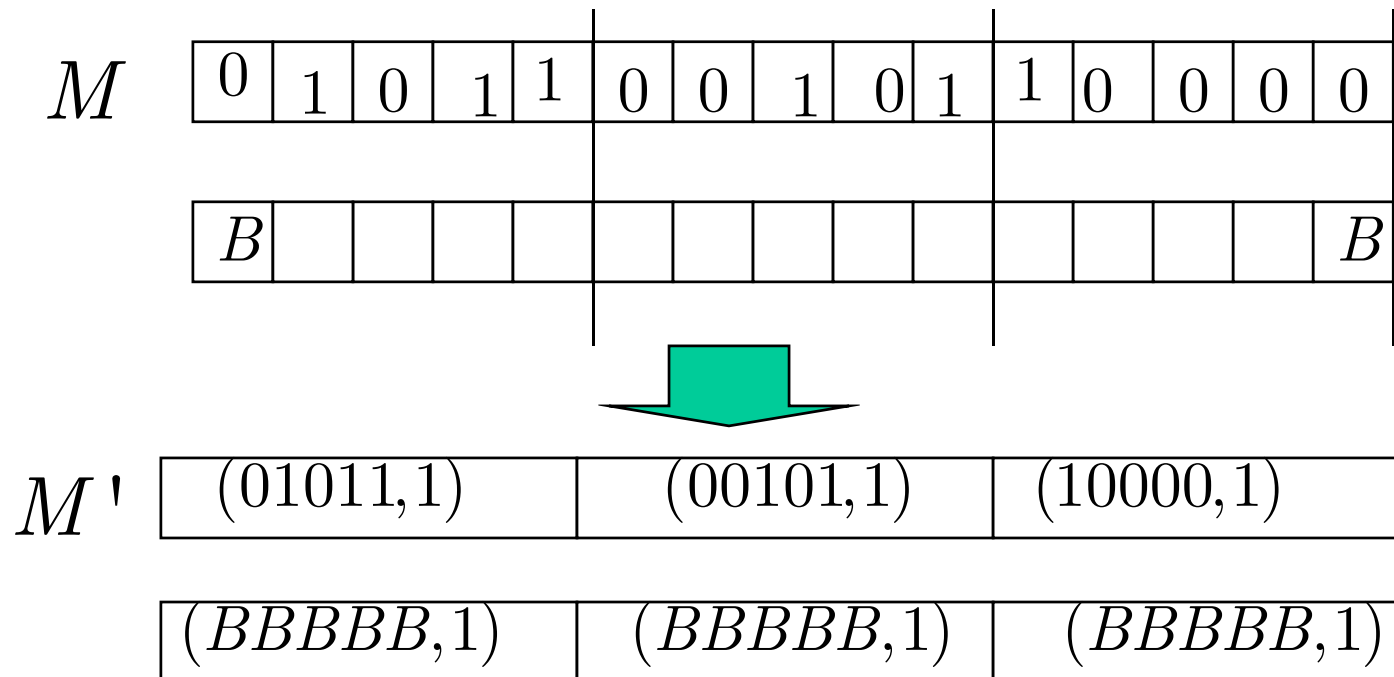
$$M \longrightarrow M' \longrightarrow F$$

と2段階で高速なTMを構成する。
 F が望みのTMである。

アイデア

Mの複数ステップをFでは1ステップで行うようにする。
そのために、Mの複数セルをFでは1セルで表現する。

簡単のために、5セル分で動作原理を調べる。
(これが一般の値でも可能であることは、
容易に確認可能である。)



- (1) 先ず、最初に与えられた0と1の入力列に対して、それを第二テープにコピーする。
ただし、5個の記号をまとめて一つの記号

$$(a_1 a_2 a_3 a_4 a_5, b)$$

とする。

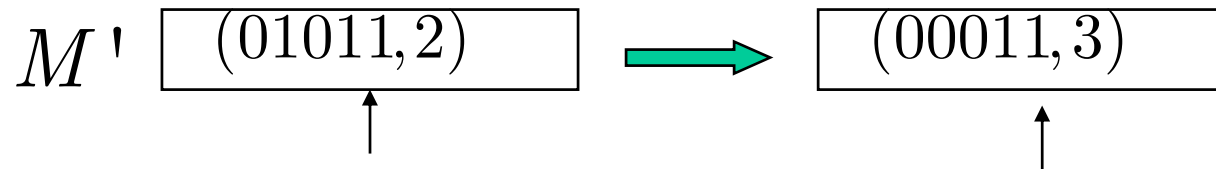
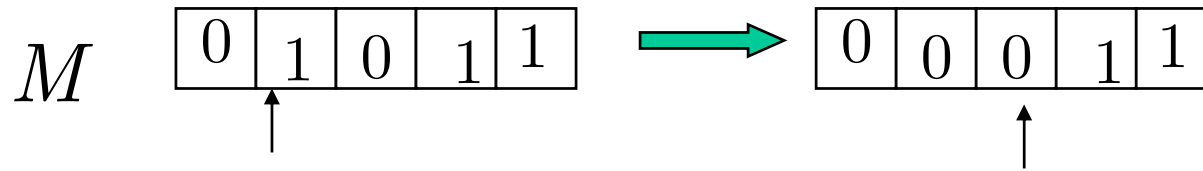
ここで、 $a_1 a_2 a_3 a_4 a_5$ はMのテープ内容で、 b はヘッド位置を表す。

$b = 1$ なら左端より左にヘッドがあり、

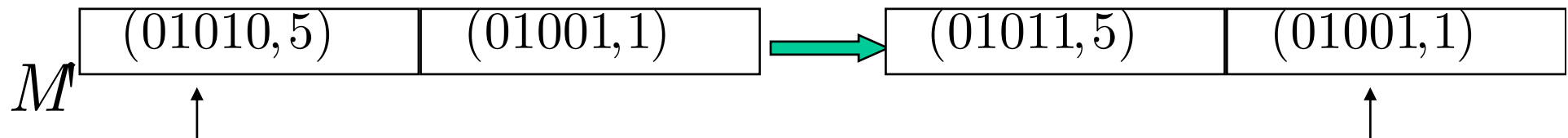
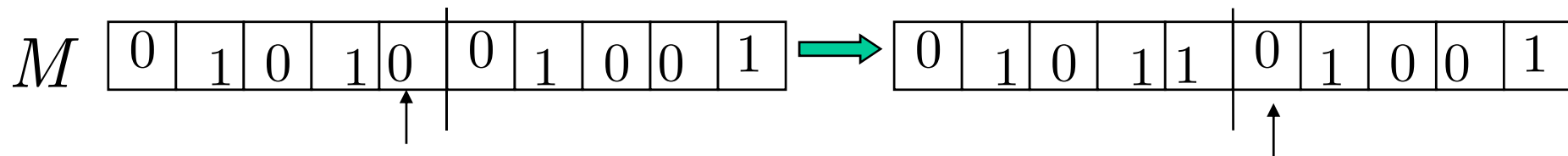
$b = 5$ なら右端より右にヘッドがる。

それらの間なら、対応する場所にヘッドがある。

(2) M' は M のシミュレーションをおこなう。



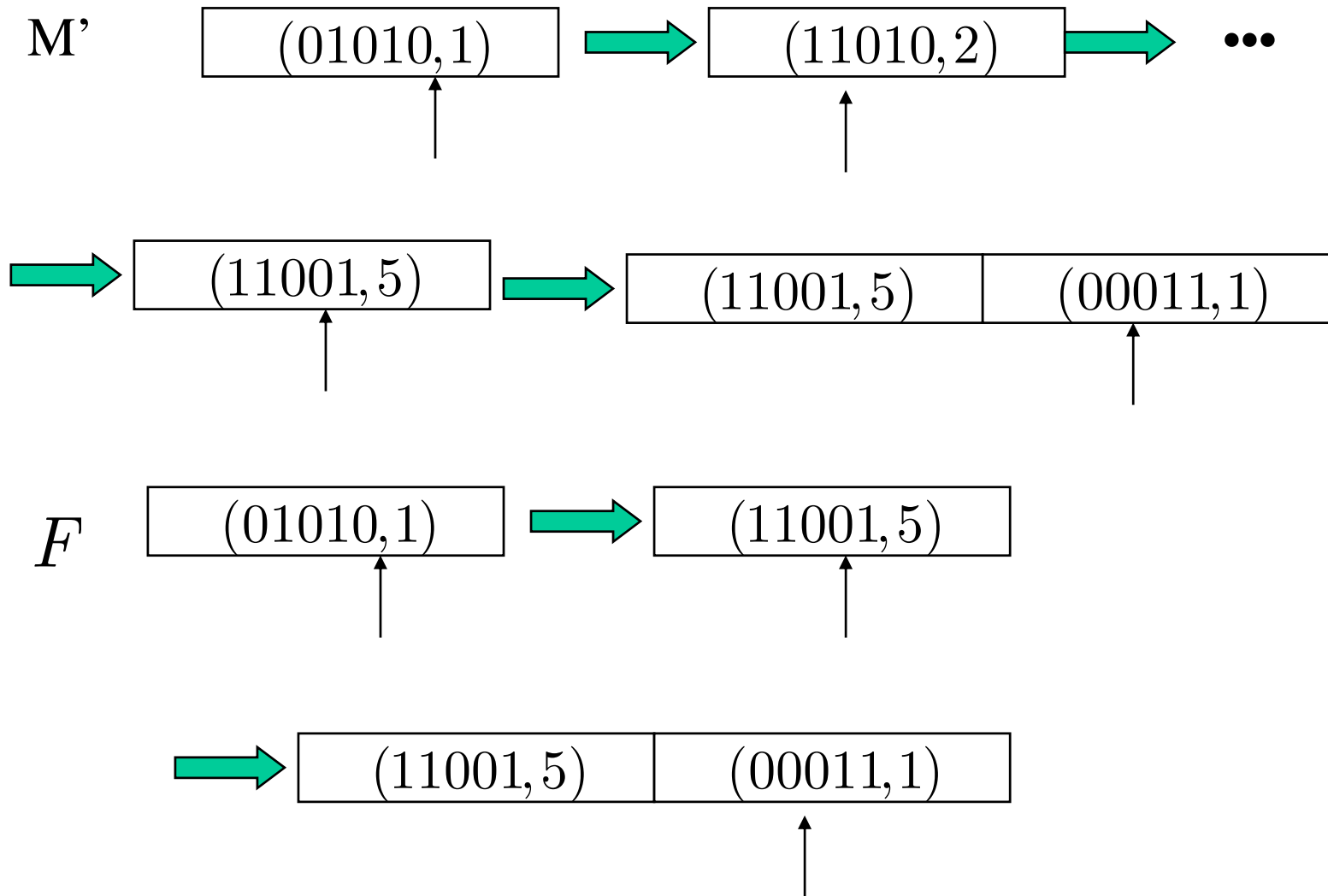
このように、 M はヘッドのヘッドの移動を、
 M' は記号の書き換えで対応させる。



M'は明らかにMの動作をシミュレートできる。
Mが受理状態になるとき、そのときに限って
M'が受理状態になるようである。

(3) 高速化

M'の動作で重要なのは、
 $b = 1$ と $b = 5$ の2つの時である。
その他の動作は捨てて($b = 2, 3, 4$ は捨てる)、
 $b = 1$ と $b = 5$ の時の動作だけをシミュレートするように
Fを構成できる。



このように、 F は M' の状態遷移を高速にシミュレートすることができる。一般に、定数の範囲での高速化が可能である。

7-4. クラスP

ここまでで、時間限定TMでは、定数倍の差はほとんど意味をなさないことをみてきた。

また、kテープTMと1テープTMとの時間の差もみてきた。

しかし、kテープTMと1テープTMの間には、多項式時間の差しかない。次の命題が成り立つ。

性質

$T(n) \geq n$ とする。

このとき、ある問題が $T(n)$ 時間限定kテープTMで解けるなら、 $O(T(n)^2)$ 時間限定1テープTMで解ける。

すなわち、多項式時間という制約においては、kテープTMと1テープTMの間に能力の差は無い。

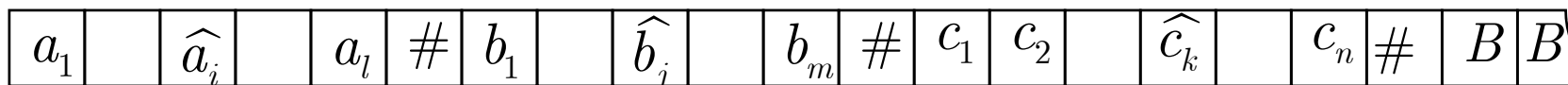
証明

方針:

M を $T(n)$ 時間限定 k テープ TM として、
 M をシミュレートする $O(T(n)^2)$ 時間限定 1 テープ TM S を構成する。

M の動作をシミュレートする 1 テープ TM S は構成できる。

S のテープの様子を示す。



実は、この S のシミュレーションが望みの時間で行えることを示す。

Mの各々のステップに対して、Sはテープの使われている部分を2回すべて走査する。

走査1:

次の動作を決定するための必要な情報を獲得する。

走査2:

走査1で獲得した情報で、動作を実現する。

Sのテープの長さは、Sの計算時間に影響するので、その長さを解析する。

まず、Mの各テープの長さを考察する。

Mは、 $T(n)$ 時間で実行されるので、

各テープの長さは、 $T(n)$ 以下でなければならない。

($T(n)$ 以上なら移動だけで $T(n)$ 時間かかってしまう。)

これより、Sのテープの長さは、

$$k \times T(n) = O(T(n))$$

である。

Mの一回のステップに対しては、2回Sを走査し、
高々k回のシフト操作を行うだけである。

したがって、Mの1ステップをSでは、 $O(T(n))$ 時間でシミュレート
できる。

よって、Mが $T(n)$ ステップで行う動作を、
Sは

$$T(n) \times O(T(n)) = O(T(n)^2)$$

時間で実現することができる。

QED

クラスPの定義

定義 : クラスP

$TIME(T(n)) = \{L \mid L \text{は } O(T(n)) \text{時間限定DTM}$
で判定される言語}

と定義する。

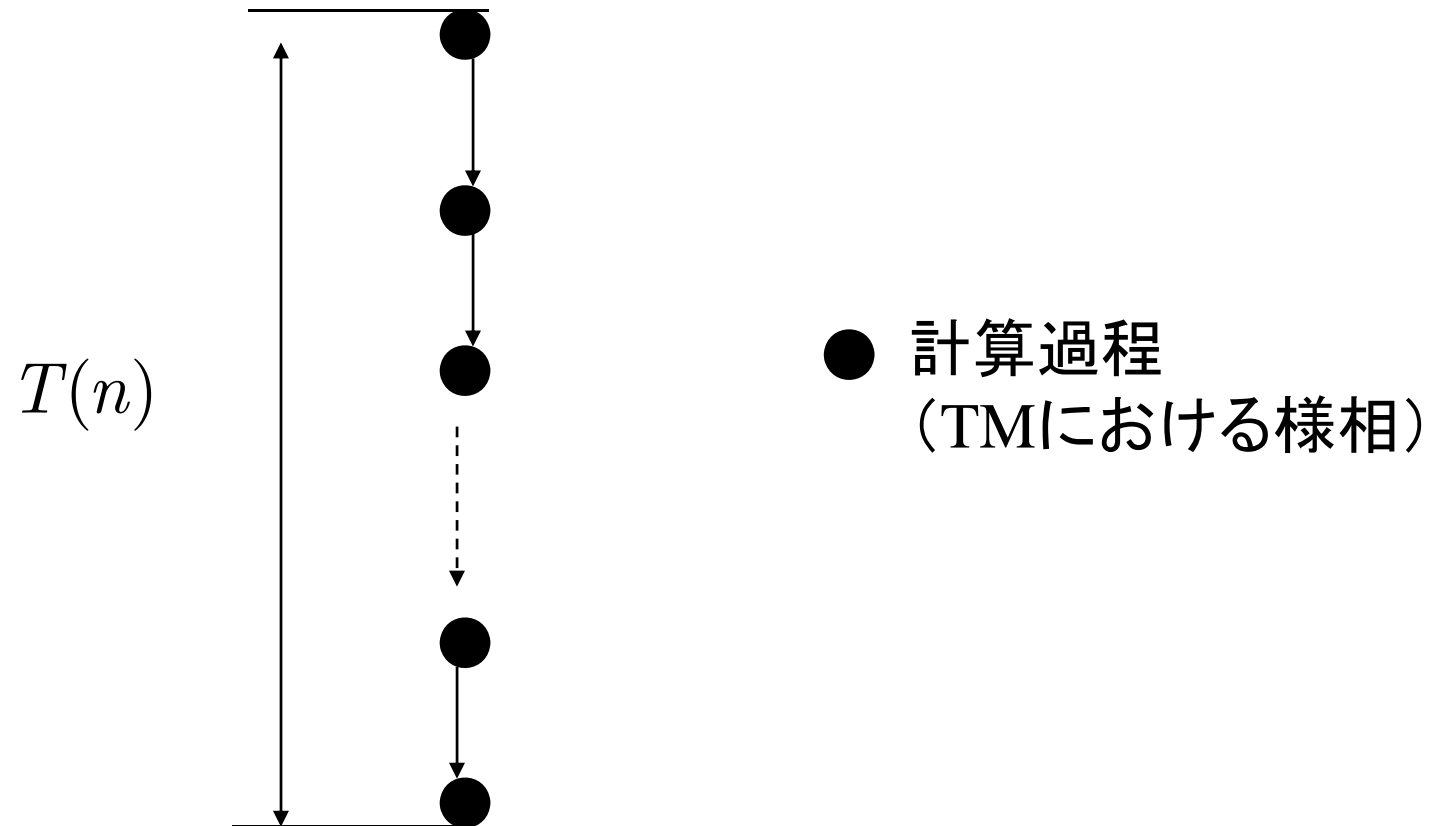
このとき、クラスPを、

$$P = \bigcup_k TIME(n^k)$$

と定義する。

つまり、通常の計算機で入力サイズの多項式時間で計算可能な問題の集合がPである。RAMでも1テープTMでも多項式時間という意味では等価となることに注意する。

決定性の計算における計算時間



決定性の計算は(入力が定まれば)一本道で行われる。
この**計算木**の高さが計算時間である。

クラスPに属する言語(問題)例

問題

名称: 回分の問題

インスタンス: 文字列 W

問: W は回文か?

すなわち、

部分文字列 x が存在して、

xx^R であるか?

問題

名称: ソート

インスタンス: 2つの実数列

$(a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n)$

問:

(b_1, b_2, \dots, b_n) は (a_1, a_2, \dots, a_n) の

置換であり、

$b_1 \leq b_2 \leq \dots \leq b_n$

か?

7-5. クラスPの意義

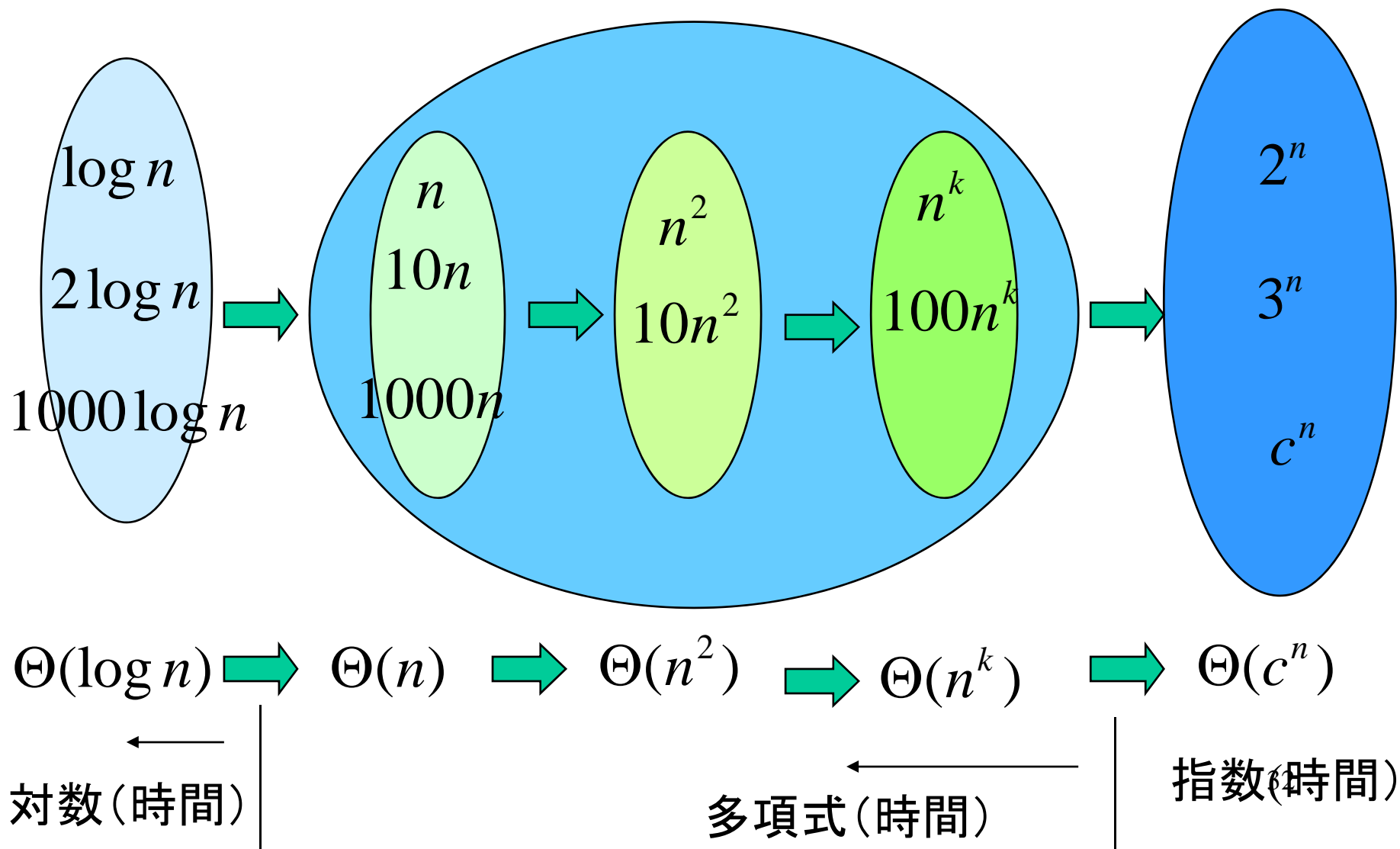
実用的に解ける問題は、クラスPに入っているとみなされる。逆の言い方をすると、クラスPに入っていない問題は、実用的な解を得ることが困難であることを意味する。

○チューリングチャーチの提言は、計算可能と不可能をチューリングマシンの存在で決めていた。

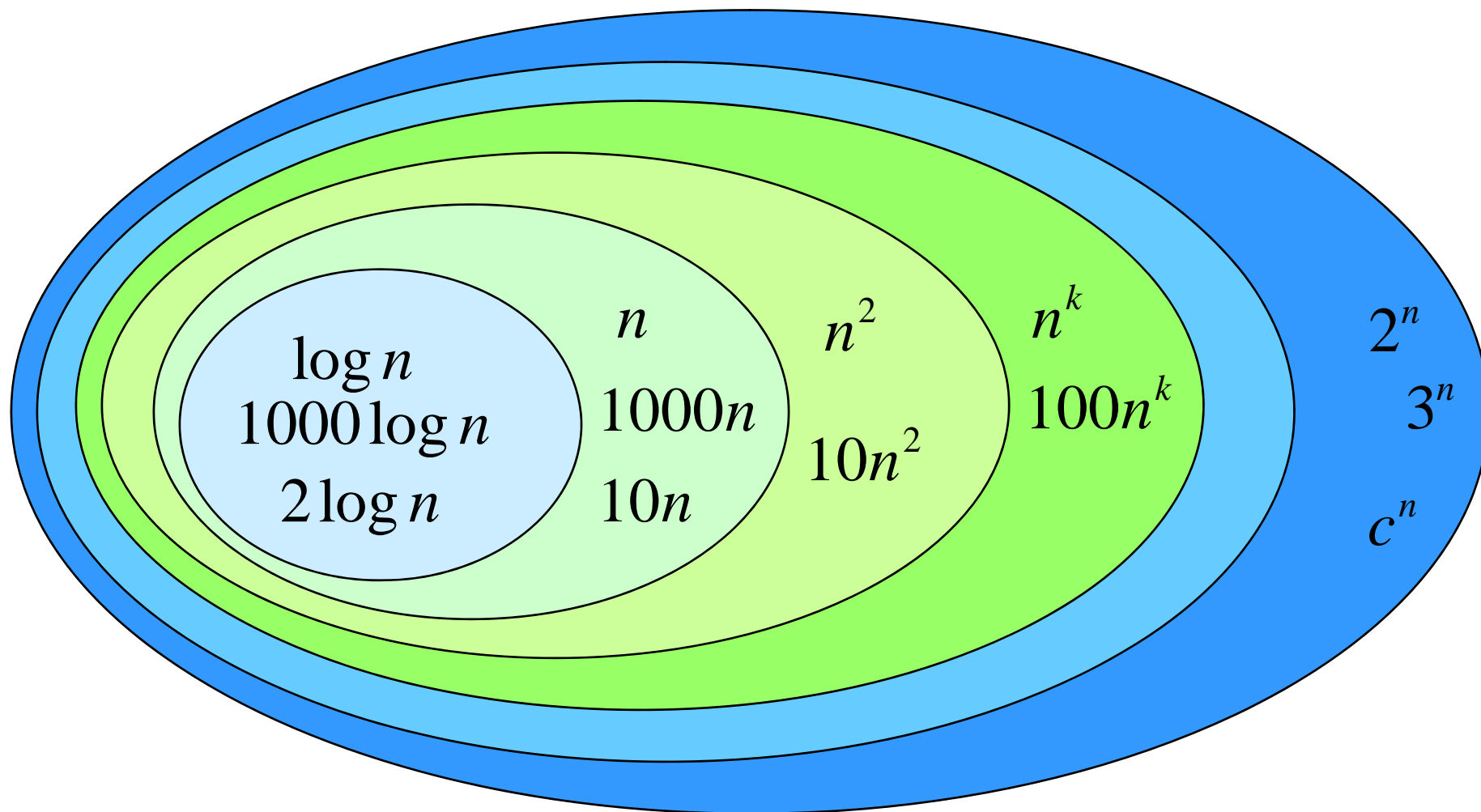
○一般に、このような有名な提言ではないが、計算機科学者間では、問題が実用的に解けるかどうかの判断基準を、「問題がクラスPに属するか」、つまり「多項式時間限定TMで解けるか？」とみなすことが多い。

ここでは、その根拠を示す。

関数の分類(シータ記法)



関数の分類(オーダー記法)



$O(\log n) \Rightarrow O(n) \Rightarrow O(n^2) \Rightarrow O(n^k) \Rightarrow O(c^n)$

←
対数(時間)

←
多項式(時間)

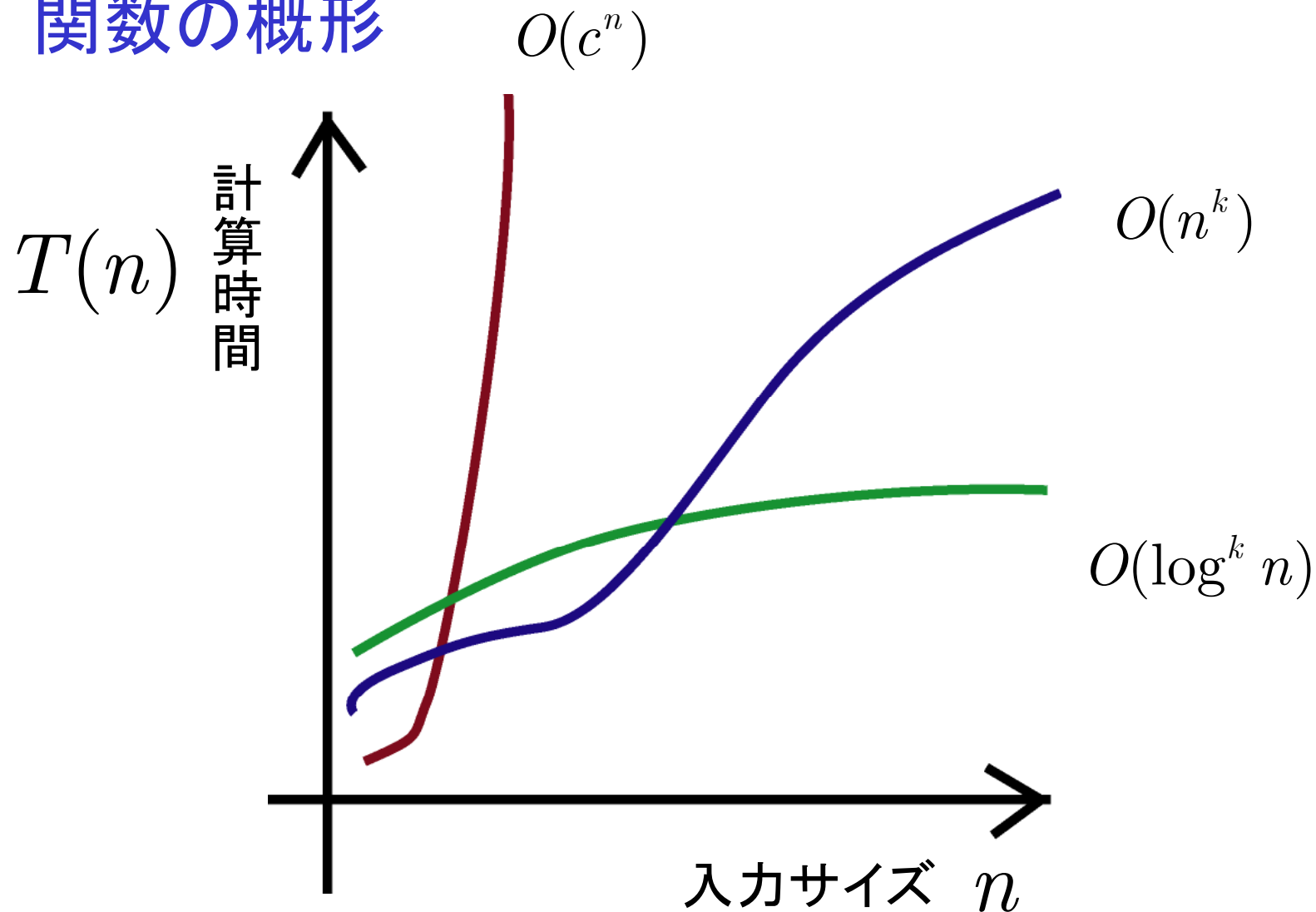
←
指数(時間)

計算時間と関数

1000MIPS(1秒間に10億回の演算可能)の
コンピュータで考えてみましょう。

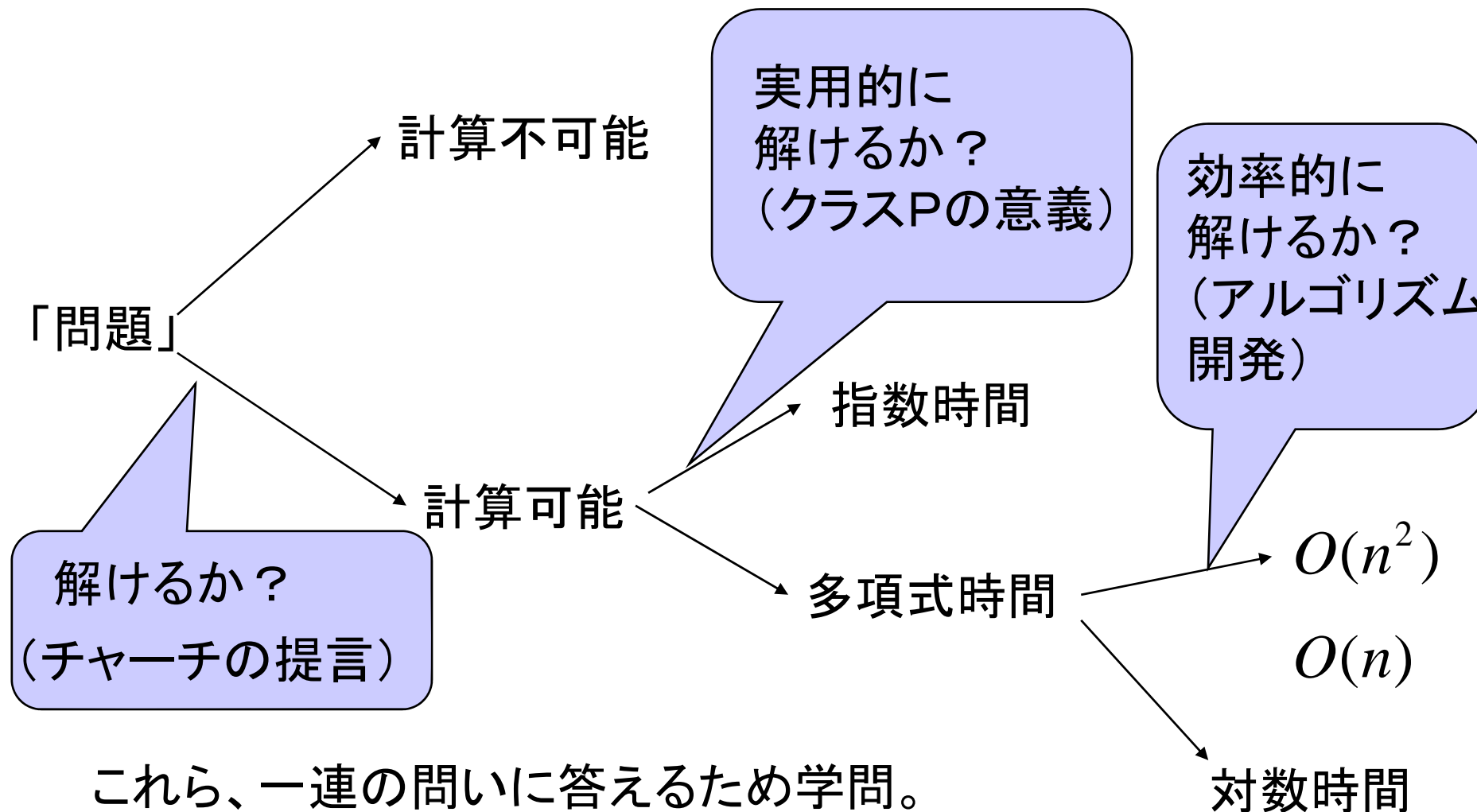
関数 サイズ n	$10^{10} \log n$	$10^8 n$	$10^5 n^2$	2^n
10	10秒	1秒	0.01秒	約 1.0×10^{-6} 秒
100	20秒	10秒	1秒	約3京世紀
10^3	30秒	1分40秒	1分40秒	甚大
10^4	40秒	16分40秒	約2時間47分	甚大
10^5	50秒	約2時間47分	約11.5日	甚大
10^6	1分	約1日	約3.2年	甚大

関数の概形



漸近的評価がいいアルゴリズムは、
多量のデータを扱うときに力を発揮する。

計算機科学の目的



これら、一連の問いに答えるため学問。
さらに、例えば同じ $O(n)$ でも、より係数を小さくすることも目的になる。