「世界」をとらえる モデル化と理論とコンピュータ

実証的な技術とともにコンピュータ(計算機)の原理や理論に力を入れている情報科学科では、「情報論理」「離散数学」「計算量理論」「形式言語理論」の理論の授業があります。理論というと、カチカチで難しい印象を受けるかもしれませんが、実はいろいろな領域と結びつく広がりがあります。ここでは、その面白さを紹介しましょう。

たいせつなのは、モノの考え方、見方、とらえる力

通常、学部の3年生で論理学を勉強する機会はなかなかありませんが、情報科学科では3年生に「情報論理」で論理学の基礎を学びます。そのねらいは「ものごと」をとらえる——モデル化、抽象化を学ぶことです。

計算機が扱えるのは数や記号だけ。計算機が本質的にしていることは記号に対する規則的な操作です。このため、論理学が情報科学の素養としてたいせつになってきます。

論理学には、構文的な側面と意味的な側面とがあります。

まず、「ものごと」を計算機で扱うためには、「ものごと」をなんらかの記号に対する操作として表現することになります。これが「形式化」で、情報論理学の授業では「ものごと」を記号や論理式というかたちでとらえます(→コラム「すべての人が誰かを愛している」)。これが論理学でいう、論理式の構文を定めて推論の規則を与える「構文論」です。

それとは逆に、形式化されたもの——記号や規則によって操作されたあとの計算結果が現実世界でどのような意味を持つかが「意味論」です。

授業では、「ものごと」を論理式というかたちでとらえる一方、論理式がどのような意味を持っているかを意味論としてとらえ、両者の関係がどうなっているかを学びます。

ものごとをどうとらえるかは、元来、系統的に勉強するのが難しいことです。けれども、記号の世界を扱い、その意味を考えることを通して、ものごとを抽象化してとらえる訓練を積むことができます。

計算をモデル化する——計算機の源に遡る

60年前にいまの計算機の原型となるプログラム内蔵式計算機ができたとき、設計者に「モデル」という明確な認識はなく、形式化も行われていませんでした。しかし、そのときにできた計算機には、「計算」とは何かということを新しい視点でとらえたモデルが内在していました。メモリに格納された「プログラム」に従って演算器で「計算」し、結果をメモリに格納するというモデル(手法)は、歯車機械で加減乗算をしていたそれまでの手回し計算機から、格段に飛躍したものだったのです。

プログラム内蔵方式の源は、計算機が生まれるしばらく前にチューリングという数学者が考えた、チューリング機械という仮想機械に遡ることができます。ほかにもこれまで、計算をするためのいろいろなモデルが考えられてきました。たとえば、関数の定義と実行を抽象化したラムダ計算というものをご存じの方もあるでしょう。

自然は計算している——いまあるものばかりが計算機ではない

計算のモデル化を考えていくと、計算機の姿がいまある計算機だけに限らないことに気付きます。

計算や情報の視点から自然をとらえると、自然現象のなかでは、情報がどんどん流れて処理されたり蓄えられたりしています。

実は生物の細胞も、外部からの多様な刺激に反応していろいろな情報処理を行っています。細胞の集まりである生物は、ひとつの超並列計算システムと見ることができます。

情報科学と生命科学が融合して生まれたバイオインフォマティクスやシステム生物学の分野では、細胞内の反応をモデル化して予想する研究が盛んになっています。たとえば、iPS細胞はどんな細胞にも分化できるもので、適切な薬をふりかけて神経にしたり肝臓にしたりすることができます。そこで、細胞中の分化の機構をモデル化してシミュレーションを行い、最終的には、細胞の分化を自由に制御することが研究の目標となっています。

また、量子力学や量子現象のなかに、情報のたたみこみや並列計算の側面を見出すことができます。この量子現象をうまく利用して、いままでできなかったような計算を可能にしようというのが、量子計算機です。

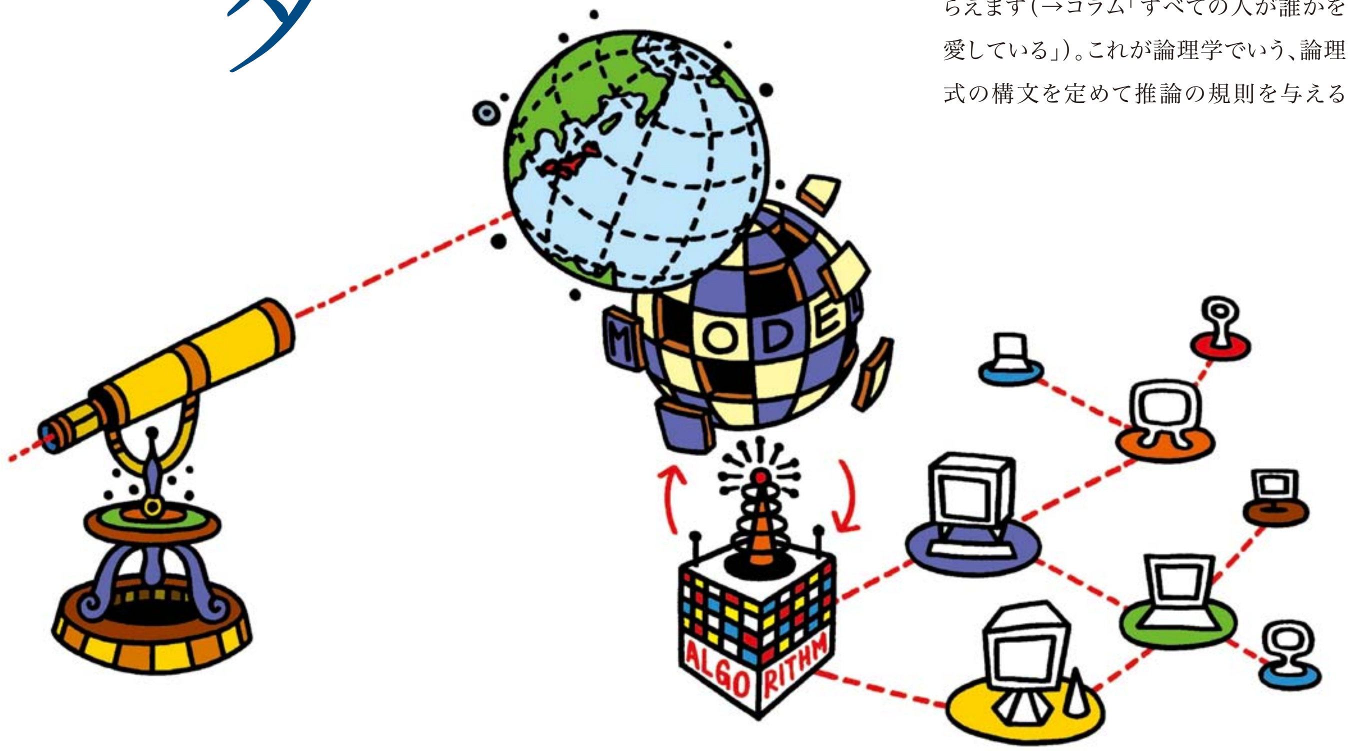
100万台の計算機を使って並列計算をする

いきなり未来の計算機の話をしましたが、計算機を大量にかき集めて並列計算をするという現在の応用シーンでも、モデル化が活きています。並列計算では、いろいろな仕事をたくさんの計算機に割り当てて計算させますが、そもそも仕事をどの計算機に割り当てるかが難問です。候補となる仕事の割り振り方はあっという間に天文学的数字になってしまうので、どの方法が良いかを真面目に比較すればたいへんな時間がかかります。計算を速くするために並列計算をするのに、仕事の割り振りに時間がかかっていては意味がありません。

そこで本質は壊さずに、数学の世界で容易に扱えるようモデル化することがポイントになってきます。モデル化を経て検証しておけば、ここまではうまくいくと保証付きです。

めぐる抽象化——記号化した世界が現実世界になって、再び抽象化の対象になる

いったん抽象化してできた記号の世界もまた、計算機で扱ったとたんに現実世界と



すべての人が誰かを愛している

Column 1

標題の文を論理式で表すと、次のようになります。

$\forall x(\exists y\ Loves(x,y))$

(任意の人xに対して、ある人yが存在し、xがyを愛している)

でも、ことによったら次のような意味かもしれません。

$\exists y(\forall x\ Loves(x,y))$

(ある人yが存在し、任意の人xに対して、x がyを愛している)

自然言語はあいまいです。普通に考えると前者になりそうですが、計算機は融通がきかないので、この2つをきちんと区別してあげる必要があります。これが形式化です。

ものごとが、いったん論理式で表されると、さまざまな推論規則を適応してあげることができるようになります。

なり、解析とモデル化の対象になります。暗号をモデル化したとき、暗号に対する攻撃者をどのようにモデル化するかということも、そのひとつです。

暗号は、計算時間を無限に使えば必ず破れます。攻撃者は懸命に計算機を回して暗号を破ろうとするので、現実世界の暗号に対する攻撃者をモデル化するときには、計算時間を限定し、暗号の強さを評価します。また、あてずっぽうでパスワードを当ててしまうこともありますから、その確率がどれくらいかを保証しなくてはいけません。そこでモデル化に、これまで扱ってこなかった計算量と確率の理論も取り込んでいくことになります。モデル化という観点から見ると、暗号は進歩している、とても難しくて魅力的な分野のひとつです。最近は、人がゲーム理論的に効率よく動くときには安全、などという理論も出てきました。

世の中が変化すると、いままでのモデル化の範疇ではできないことが現れます。それを埋めるようにモデルを拡張していけるところが理論家の醍醐味です。

ばらばらの世界を取り扱う——計算機ができておもしろくなった離散数学

「離散数学」は、「離散」つまり「ばらばらなもの」を対象とし、その関係や組合せなど、問題が本質的に持つ構造をとらえるための数学です。離散数学は、計算機ができて実際に計算が可能になり、アルゴリズムが重要になって、広がりがでてきた分野です。

AさんBさんとCさんの3人がいるとき、AさんとBさんは仲が良いけど、Cさんが2人を嫌いだったら、Cさんは孤立しています。そういう関係性を表すのがグラフ理論です。

これに対して、いまの時間はここ、何かが来るか来ないかで次の時間にはそこかあそこ、その次にはさらに……と状態の遷移を表現するのがオートマトンです(これは「形式言語理論」で勉強します)。

現実の世界に目を向けると、Webページのリンクはグラフ構造を持っていますし、ネットワーク通信の制御は状態遷移図で表せます。このような実世界のさまざまな現象や関係を、離散数学で扱うことができます。

さて、計算機で何かしようとしたとき、データにいい構造があるということは、いい方式(アルゴリズム)があるということ。いい構造がないと計算は難しく、実用性がないほど時間がかかります。どの程度難しいかをランク付けし、どうしようもないものはスジが悪いとレッテルを貼るのが計算量理論です。

「計算量理論」の授業では、チューリング機械という計算モデルに始まって、最先端の計算量理論までを解説します。

理論家は月に向かって刀を研ぐのか?

情報科学科/コンピュータ科学専攻では、100万台の計算機を使って並列計算をする、あるいは暗号システムを作るというような、いまの計算機を題材にした問題を解くとともに、生物分子計算、量子計算といった次世代の計算機を題材にしたモデル化を手がけています。

しかし、モデル化を中心とした理論に対する風当りは、ときとして強いことがあります。たとえば、形式手法の研究者に対して「月に向かって刀を研ぐ」という悪口があります。結果を急いで手に入れたい実践家にとって、モデル化の手法の開発から始めることは、結果からとても遠いところにいるように思えるのかもしれません。

しかし、実践的な研究、つまりいまあるものをしゃぶりつくすような手法の開発の多くは、10年もたてば廃れてしまうことでしょう。これに対して、理論から出た結論はすぐに使われなくても、50年後、100年後まで残っていきます。

言い換えると、理論はその枠組みのなかでは、決して廃れることなく、永遠に生き続け、100年後に突然役に立ったりするわけです。ですから、将来、月を切る刀が本当にできるかもしれません。

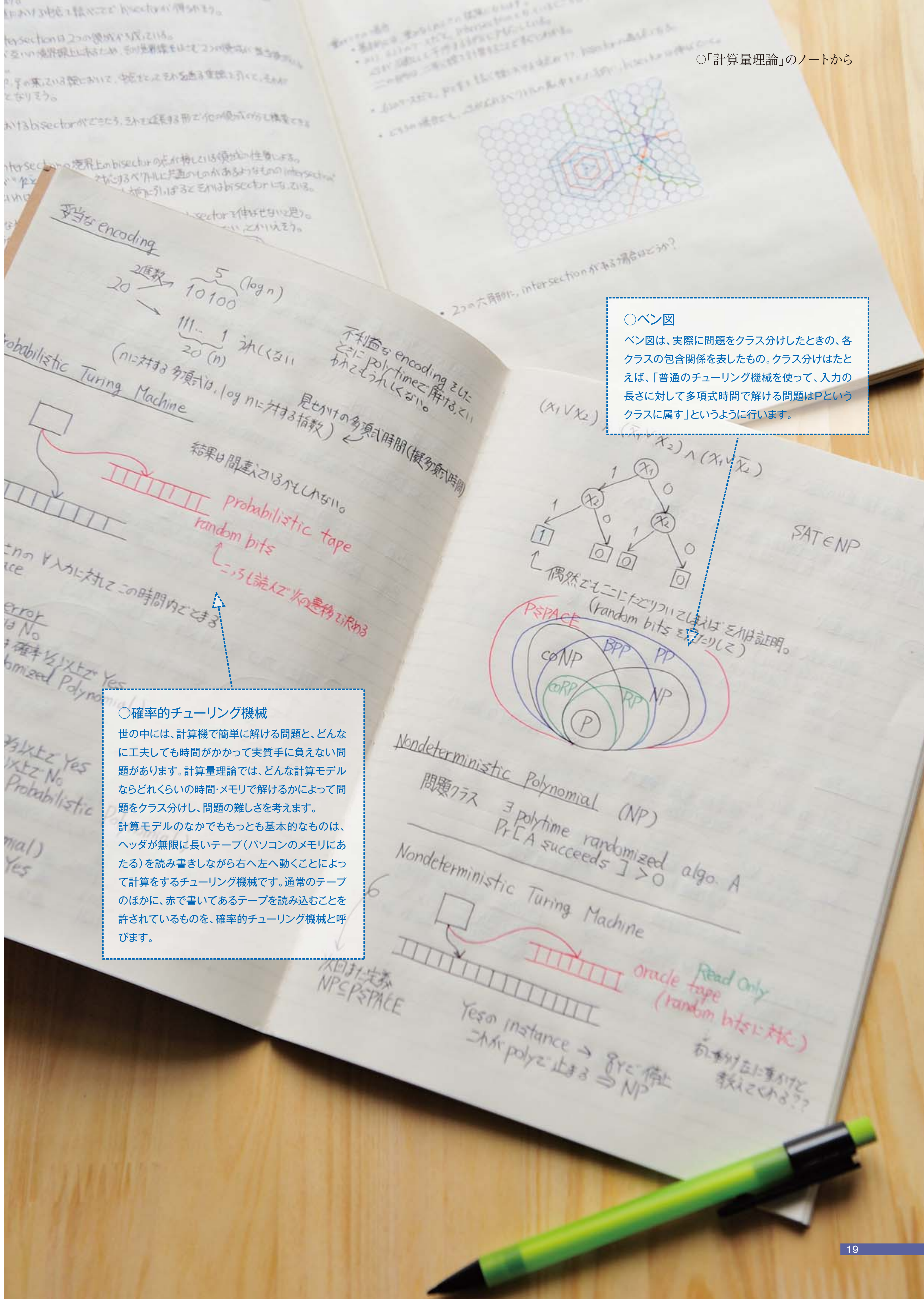
もちろん、いまの世界と接点を持ち、すぐにも役に立つような理論もありますが、理論の研究の最前線では、やっぱり新しいことをやっていくという、開拓地にいる感じがあります。そのようななかで、50年、100年先のことをやれるのは、まさに理論屋の冥利でしょう。

情報を失う計算はエネルギーを生じる　Column 2

1970年代は、「計算の物理」が盛んに研究されていました。その中心には、「情報を失う計算はエントロピー的にエネルギーを発生する」というランダウアーの原理があります。たとえば、回路のAND素子は2つの入力に対して出力が1つなので、そこで1ビットの情報が失われます。この失われた情報分のエネルギーが消費されてしまうので「計算にはエネルギーが必要になる」というのです。

そこで、現在では量子計算で有名なベネットは、情報を失わない可逆計算ならエネルギーを使わずにできるのではないかと考え、DNAの転写反応をモデルにした仮想的な分子機械を考えました。平衡状態で可逆的に動けば、ほとんどゼロのエネルギーで計算できるというものです。

その後、エーデルマンが実際にDNAで計算可能なことを示し、並列計算の観点とともに消費エネルギーも注目を集め、生物分子計算の端緒となりました。



○「計算量理論」のノートから

○ベン図
ベン図は、実際に問題をクラス分けしたときの、各クラスの包含関係を表したもの。クラス分けはたとえば、「普通のチューリング機械を使って、入力の長さに対して多項式時間で解ける問題はPというクラスに属す」というように行います。

○確率的チューリング機械
世の中には、計算機で簡単に解ける問題と、どんなに工夫しても時間がかかって実質手に負えない問題があります。計算量理論では、どんな計算モデルならどれくらいの時間・メモリで解けるかによって問題をクラス分けし、問題の難しさを考えます。
計算モデルのなかでももっとも基本的なものは、ヘッダが無限に長いテープ(パソコンのメモリにあたる)を読み書きしながら右へ左へ動くことによって計算をするチューリング機械です。通常のテープのほかに、赤で書いてあるテープを読み込むことを許されているものを、確率的チューリング機械と呼びます。