

物理シミュレーションと時系列データからの事象認識による 物理問題解答

Physics Problem Solving Based on Physics Simulation
and Event Recognition from Time-Series Data

横野 光*¹ 稲邑 哲也*¹
Hikaru Yokono Tetsunari Inamura

*¹国立情報学研究所
National Institute of Informatics

We tackle to a system to solve physics tests of university entrance examination by integrating natural language processing and physics simulation. Our system first understands a situation described in a question and simulates the situation. Then, it interprets the result and return answer. In this paper, we discuss components and resources of our system.

1. はじめに

与えられた状況を理解し、そこから先に起こる事象を予測し、それに基づいて次の動作を決定するという事は、我々人間が日常的に行っている活動であり、より柔軟な対話システムの構築において重要な要素であると考えられる。しかし、システムの構築において、与えられる状況とはどういうものか、何を予測しなければいけないのか、どのように予測すれば良いのか、など考えなければならないことは非常に多い。

我々の目的はそのような状況理解と将来予測のフレームワークの構築であるが、まずその応用の一つとして大学入学試験の物理問題解答というタスクを考え、それに対するシステムの構築を通して解くべき課題を明確にし、必要となる処理や資源についての知見を得ることを目指す。

物理の問題は物理現象の理解を問うものであり、出題される状況は物理現象に関するものだけに焦点が当てられている。従って、本タスクにおいて理解すべき状況は物理的な世界に関するものに限定できるため、予測すべき事象や状況を定義することは困難ではない。

また、問題では初期の状態からどのような物理現象が発生するかが書かれていることが多い。提案システムでは初期の状態を理解し、そこから発生する事象を予測する。そして、その結果に基づいて問題に解答する。このとき、予測すべき事象は発生する物理現象であり、どのようなことが起こるかは物理シミュレータによって予測することが可能である。提案システムの核となる要素は、1) 問題に書かれている状況の理解、2) シミュレーション結果の解釈 である。

本稿では、提案システムの概要を述べ、必要となる資源と処理について議論する。

2. 物理問題の特徴

大学入学試験の物理科目において、多くの問題では図1に示すように問題の対象となる状況についての説明がなされ、その状況から発生する物理現象に対してそのときの物理量に関する問いが出題される。

問題文は受験者が読み間違いをしないように統制されて記述されており、一般的な文章に比べて容易に理解することができる。

連絡先: 横野光, 国立情報学研究所, 〒101-8430 東京都千代田区一ツ橋 2-1-2, yokono@nii.ac.jp

問4 あらひ水平面上での質量 m の物体の運動を考える。図3のように、一端を固定したばねを自然の長さから d だけ縮め、他端に物体を置いて手で押さえた。次に、手を離すと物体は初速度0で動き始め、やがてばねから離れて、手を離れたときの位置から x だけ進んで止まった。 x として正しいものを、下の①～⑥のうちから一つ選べ。ただし、ばねの質量は無視できるものとする。また、ばね定数を k 、面と物体の間の摩擦係数を μ' 、重力加速度の大きさを g とする。

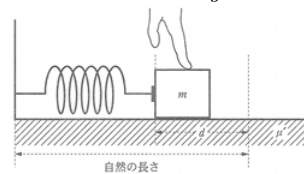


図3

- ① $\frac{k}{2\mu' mg} d^2 + d$ ② $\frac{k}{\mu' mg} d^2 + d$ ③ $\frac{k}{2\mu' mg} d^2$
④ $\frac{k}{\mu' mg} d^2$ ⑤ $\frac{2\mu' mg}{k}$ ⑥ $\frac{\mu' mg}{k}$

図1: 物理の試験問題の例 (2009年度センター試験物理追試験第1問より引用)

る。しかし、それは必ずしも計算機にとっても容易であるということの意味しない。問題文では曖昧性が生じない程度に省略などが使用されることがある。例えば、以下の文章を考える。

- (1) ばねの一端を天井に取り付け、他端におもりを取り付ける。そして、おもりをばねの長さが d になるまで引っ張った後で離れた。おもりは振動しやがて静止した。

ここで“離れた”には2つの解釈が考えられる。

- (2) おもりを 手 から離れた
(2') おもりを ばね から離れた

我々は(2)の解釈が正しいと判断できるが、このようなゼロ代名詞の解析は計算機に困難なタスクの一つである [Iida 11]。

人間はこのような曖昧性の解消に常識を用いていると考えられるが、計算機に常識を用いた推論を行わせることは困難である。しかし、本タスクではドメインを大学入試の物理の問題に限定しているため、必要な常識を限定することができる。従って、必要な知識を構築し、それを使った推論は可能であると考えられる。

3. シミュレーションを用いた物理問題解答

人間が実際に問題を解くときには、まず問題となっている状況を理解し、その背後にある物理現象を推測する。そして、その物理現象に係る公式を用いて答えを導く。これに対して、本研究で構築するシステムでは、問題となっている状況を理解し、物理シミュレータによってその先で起こる事象のシミュレーションを行い、その結果を使って問題に解答する。

システムの流れを図2に示す。システムは自然言語で書か

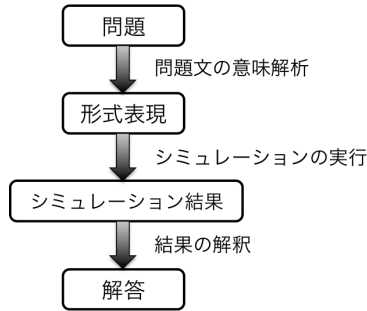


図2: システムの流れ

れた問題に対して、そこに書かれてある状況を形式表現に変換し、それをもとにしてシミュレーションを行う。そして、シミュレーション結果と問題の形式表現から問題として問われている時点を同定し、その時点における該当の物理量の値を用いて解答する。問題には状況の説明を図で表現している場合があるが、現時点では図の情報は利用せず、テキストのみを考慮する。

3.1 問題文の意味解析

自然文で記述された問題からそこに書かれている内容を理解する。物理の試験では、記述された状況において起きた物理現象に関する理解が問われる。そのため、必要となる情報は物理現象に関わるものだけに限定することができ、その他の要素は無視できる。例えば、“探査機を惑星に着地させる”や“ボールを床に落とす”という状況を考えたとき、“探査機”や“ボール”がどういうものか、ということを考える必要はなく、これら両方の事象が“物体が落下する”という物理現象における“物体”であることが理解できればよい。従って、定義すべき述語の数は限定できる。

本研究では、a) 物体、b) 物体の属性、c) 操作、d) 物理現象の4種類に対する述語を定義する。表1に述語の一部を示す。物体に対する操作と物理現象は時間軸に沿って順番に発生する。これらを意味する述語に対して、イベント変数 e_i [Davidson 80] を導入し、この変数の順序関係 ($e_i < e_j$ iff $i < j$) によって事象間の時間的な順序関係を表す。例えば、“車が出発して、点 p に到着した”という文は以下のように表すことができる。

$$car(c) \wedge point(p) \wedge start(c, e_1) \wedge reach(c, p, e_2)$$

実際に問題に解答するためには、問題で問われている値が何かを同定する必要がある。本研究における形式表現では、この問われている値に対応する変数を存在量化して明示的に表す。図1の問題に対する形式表現を図3に示す。

形式表現への変換では、システムは例えば“水平面”が述語 $floor$ で表されるものと同定しなければならない。この変換の一つの方法として、単語とそれに対応する述語を記述した辞書

表1: 述語の例

	述語	説明
a) 要素	$floor(x)$	x は床である
	$mass(x)$	x は物体である
	$fixpoint(x)$	x は視点である
	$spring(x)$	x はばねである
	$hand(x)$	x は手である
b) 属性	$weight(x, m)$	x の質量が m である
	$cnst_spring(x, k)$	x のばね定数が k である
	$distance(x, y, d, e_i)$	e_i における x と y 間の距離が d である
	$dfc(x, y, u)$	x と y 間の動摩擦係数が u である
c) 操作	$ga(g)$	重力加速度が g である
	$connect(x, y, e_i)$	x と y が e_i で接続される
	$shorten(x, d, e_i)$	e_i で x が d だけ縮められる
	$touch(x, y, e_i)$	e_i で x は y に接触する
	$hold(x, y, e_i)$	e_i で y が x を把持する
d) 物理現象	$release(x, y, e_i)$	e_i で y が x を離す
	$stop(x, e_i)$	e_i で x が静止する
	$move(x, e_i)$	e_i で x が動く

$$\exists x(floor(o) \wedge mass(a) \wedge weight(a, m) \wedge fixpoint(p) \wedge spring(s) \wedge cnst_spring(s, k) \wedge connect(p, ep1(s), e_1) \wedge hand(h) \wedge shorten(s, d, e_1) \wedge distance(a, ep2(s), 0, e_1) \wedge hold(a, h, e_1) \wedge release(a, h, e_2) \wedge move(a, e_2) \wedge stop(a, e_3) \wedge ga(g) \wedge dfc(o, a, u) \wedge distance(a, ep2(s), x, e_3))$$

図3: 図1の問題の形式表現

を用意するということが考えられる。しかし、以下の文における“レンガ”はそれぞれの文中で異なる役割を持つ。

(3) ボールをレンガに落とす。

(3') レンガを落とす。

文(3)では“レンガ”は床と同じ扱いであるため述語 $floor$ で表されるが、文(3')では $mass$ という述語で表される。このように実際の変換では文脈を考慮を考慮する必要がある。

3.2 物理シミュレーション

3.1節で述べた形式表現を基にして、物理シミュレーションを行う。本研究では物理シミュレータに OpenModelica*1 を採用する。このシミュレータの入力形式は Modelica 言語と呼ばれるものであり、我々の定義した形式表現とは形式が異なるため、形式表現から Modelica 言語への変換を行う必要がある。この変換は基本的には述語とそれに対応する Modelica のコンポーネントの変換規則を用いて行う。現時点では対応するコンポーネントが存在しない述語があり、それらに対しては独自のコンポーネントを作成する。

シミュレーションに必要な情報は問題に記述されている状況に出現している物体と、その属性、そして、それらに対して行われる操作であり、これらに対応する述語からシミュレータへの入力を作成する。

シミュレーションの実行において、“質量 m ”などのように問題では変数として与えられている物理量に対して、適当な実数値を代入する必要がある。このとき、代入する値によっては状況が正しく再現されない可能性がある。例えば、図1の問題で、物体の質量を大きくしすぎると、手を離してもその位置から動かないという状況が起こりうる。このような問題の解決

*1 <https://www.openmodelica.org/>

策としては、3.3節で述べるシミュレーション結果の解釈と書かれてある物理現象との照合を利用するという方法が考えられる。即ち、シミュレーション結果から問題に書かれてある物理現象が発見できなければ、割り当てた実数値が不適切であると判定し、割り当てをやり直す。

3.3 シミュレーション結果の解釈

シミュレーションの結果は物理量の時系列データとして与えられる。問題ではある物理現象が起きたときの物理量の値が問われるため、シミュレーション結果から指示された物理現象が起きた時点を同定しなければならない。

3.1節で述べたように、解答すべき値は存在量化されている。問題として指示されている物理現象は、この存在量化された変数を項に持つ述語と同じイベント変数を持つ述語を探すことで同定できる。例えば、図3に示す形式表現では、存在量化されている値は x であり、それを含む述語は $distance(a, ep_2(s), x, e_3)$ である。この述語は“ e_3 における a と $ep_2(s)$ 間の距離が x である”ということであり、イベント変数 e_3 を項に持つ述語は $stop(a, e_3)$ である。この述語の意味は“ a が e_3 で停止した”ということであり、実際の問題文と対応している。

次に、同定した物理現象に対して、シミュレーション結果からその事象が起きた時点を同定する。

シミュレーション結果は物理量の時系列データなので、物理現象の意味を物理量の変化に関する条件式として定義し、条件を満たす時点でその物理現象が起きたと見なす。例えば、“止まる”という現象は対象の速度がそのときに0になり、それ以前では0より大きな値を持っていたと解釈できる。従って、“止まる”の意味は次のように形式化できる。

$$stop(obj, e_i) ::= obj.velocity(e_i) == 0 \\ \wedge obj.velocity(e_{i-1}) > 0$$

ここで、 $obj.velocity(e_i)$ は obj の時点 e_i における速度を表す。実際に我々が定義した物理現象の意味の一部を表2に示す。図1の問題に対するシミュレーション結果を表3に示す。このシミュレーションの実行に必要な値は人手で設定した*2。

表 3: シミュレーション結果

Time	x	Velocity of a
⋮	⋮	⋮
10.786	1.27535595	0.00088038
10.788	1.27535751	0.00068438
10.790	1.27535868	0.00048838
10.792	1.27535946	0.00029238
10.794	1.27535985	0.00009638
10.796	1.27535990	0.00000000
10.798	1.27535990	0.00000000
⋮	⋮	⋮

上述したように、この問題で問われているのは e_3 における x の値であり、その e_3 は a が止まった時点である。 $stop(a, e_3)$ を満たすのは表3から $t = 10.796$ の時点であることが分かり、このときの x の値は 1.27535990 である。

最後にこの値を使って問題に解答する。問題の解答形式は数値や式、グラフなど様々であり、シミュレーション結果から得られた値をその形式にあわせて最終的に出力しなければならない

ない。例えば、図1の問題では x を表す式を答えるように指示されている。このような場合は、選択肢の各式にシミュレーションで使用した値を代入し、その値とシミュレーション結果との比較を行い、その差が最も小さいものを解と見なす。

この問題の各選択肢に実際に使用した値を代入すると以下のようになる。この中で③がシミュレーション結果から導かれ

- ① 1.775510 ② 3.051020 ③ 1.275510
④ 2.551020 ⑤ 0.196000 ⑥ 0.098000

た値に最も近いものであり、これは実際の正解と一致する。

4. 評価

先に述べたように対象ドメインを物理の試験に絞ることで、考慮すべき語彙や述語を限定できる。従って必要となる言語資源を人手で構築することは不可能ではないと考えられる。十分網羅的な言語資源を構築するために必要なデータのサイズを推定するために、語彙の被覆率の調査を行った。

過去15回分のセンター試験から力学の問題のみを対象とし、状況の説明に用いられている名詞句と動詞句を人手で抽出した。また、それらに対し対応する述語を人手で付与した。定義した述語と名詞句の対応を表4に示す。動詞に関しては、同義語をまとめそれぞれに対応する述語を定義した。

表 4: 述語と対応する名詞句

述語	名詞句
$fixed_pulley(x)$	定滑車, 滑車
$line(x)$	糸, ひも, ワイヤロープ
$stand(x)$	スタンド
$stick(x)$	棒, 角棒
$hand(x)$	手
$floor(x)$	水平面, 面, 地面, 床, 板, 斜面, 地上
$groove(x)$	溝
$motor(x)$	モーター
$fixpoint(x)$	支点, ちょうつがい, 釘, 壁
$ceiling(x)$	天井
$mass(x)$	おもり, 物体, 小球, 惑星, 小鳥, 台車, エレベーター, かご, 砂, 容器, 台, 箱
$spring(x)$	ばね, ゴムひも
$pendulum(x)$	振り子

15回分の試験を学習データとテストデータに分割し、学習データに出現する単語と述語で言語資源を構築したと仮定して、テストデータに対する言語資源の被覆率を以下の式から求めた。

$$coverage(T, S) = \frac{|T \cap S|}{|T|}$$

T は学習データに出現する単語の集合、 S はテストデータに出現する単語の集合である。

評価は leave-one-out 方式の交差検定で行った。学習データは1回分の試験を使った場合から、14回分の試験を使った場合までのそれぞれに対して被覆率を求めた。学習データに追加する順番による偏りを軽減するため、追加する順番を無作為にし、各試行に対して10回計算を行い、その平均を最終的な値とした。結果を図4に示す。“単語”は単語レベルでの被覆率、“述語”は単語を述語に変換したときの被覆率を示す。

名詞に関しては、10回分の試験を学習データとして使用した時点でほぼ飽和しており、被覆率も十分に高いため定義した

*2 $m = 1, k = 1, d = 0.5, \mu' = 0.01, g = 9.8$

表 2: 物理現象に対する意味記述の例
意味記述

述語	意味記述
$fall(obj, e_i)$	$obj.position_y(e_i) < obj.position_y(e_{i-1})$
$pass(obj_1, obj_2, e_i)$	$obj_1.velocity(e_i) > 0 \wedge obj_1.position(e_i) == obj_2.position(e_i) \wedge \neg(obj_1.position(e_{i-1}) == obj_2.position(e_{i-1}))$
$leave(obj_1, obj_2, e_i)$	$obj_1.position(e_{i-1}) == obj_2.position(e_{i-1}) \wedge \neg(obj_1.position(e_i) == obj_2.position(e_i))$
$moveup(obj, e_i)$	$obj.velocity(e_i) > 0 \wedge obj.position_y(e_i) > obj.position_y(e_{i-1})$
$shorten(obj, e_i)$	$obj.length(e_i) < obj.length(e_{i-1})$

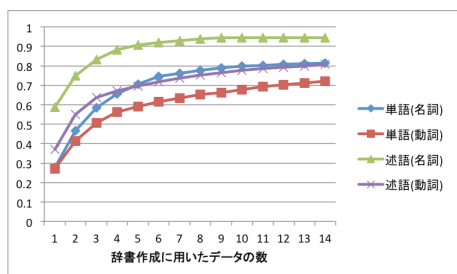


図 4: 被覆盖率

述語でほぼ網羅できると考えられる。これに対し、動詞に関しては述語レベルでも約 0.8 にとどまっている。これは問題で使われている動詞の種類が多様であるからだと考えられる。

5. 関連研究

状況認識に関しては、物理世界での状況を自然言語で表現し、質問応答を行う SHRDLU というシステムが提案されている [Winograd 87]。このシステムでは基本的な語彙と文法を与えておくことで、積み木の世界における物体の参照や曖昧性解消、質問生成などが可能であることが示されている。この他にもエージェントの視点から見た物体の位置関係を考慮した照応解析を行うシステム [Shinyama 00] などが提案されている。しかしながら、これらのシステムで対象としている世界は静的なものであり、与えられた状況からどのようなことが起こるか、までは考慮されていない。

自然言語の文から形式表現への変換には、より深い言語処理と推論が必要となる。このようなタスクは semantic parsing [Poon 09] と呼ばれ、例えば自然文で入力された命令をロボットの操作言語に変換する研究などが行われている [Liang 09]。

本研究においてシステムは物理量の時系列データからそこで発生した事象を認識する。これは数値データと言語表現間の関係を捉えるタスクとみなすことができる。このようなタスクの目的の一つはユーザに分かりやすい形で情報を提供することである。例えば、医療データから自然言語の表現を生成するという研究がなされている [Portet 09]。また、非言語情報と言語情報との対応という観点では、画像からその説明文を生成するという研究が行われている [Kuznetsova 12]。

6. おわりに

本研究では、物理シミュレーションと自然言語処理を組み合わせた物理解答システムについて述べた。既存の質問応答システムでは既にかかれてあることのみを対象としているが、本システムはシミュレータを使って状況の変化を予測し、シミュレーション結果に基づいて問題に解答する。

現時点では我々の対象は大学入試であるが、本研究は Human-Robot-Interaction にも応用できると考えている。例えば、対話でロボットに指示するとき、その指示が曖昧であれば、ロボットはユーザに問い直すだけでなく、本フレームワークを用いて適切な状況を推測することも可能になる。

また、我々の現在の焦点は問題解答であるため、自然文で書かれている状況の認識に焦点を当てているが、将来的にはシミュレーション結果を説明する文の生成にも取り組む予定である。

参考文献

- [Davidson 80] Davidson, D.: *Essays on Actions and Events*, Oxford: Clarendon Press (1980)
- [Iida 11] Iida, R. and Poesio, M.: A Cross-Lingual ILP Solution to Zero Anaphora Resolution, in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pp. 804–813 (2011)
- [Kuznetsova 12] Kuznetsova, P., Ordonez, V., Berg, A., Berg, T., and Choi, Y.: Collective Generation of Natural Image Descriptions, in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 359–368 (2012)
- [Liang 09] Liang, P., Jordan, M., and Klein, D.: Learning Semantic Correspondences with Less Supervision, in *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pp. 91–99 (2009)
- [Poon 09] Poon, H. and Domingos, P.: Unsupervised Semantic Parsing, in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pp. 1–10 (2009)
- [Portet 09] Portet, F., Reiter, E., Gatt, A., Hunter, J., Sri-pad, S., Freer, Y., and Sykes, C.: Automatic generation of textual summaries from neonatal intensive care data, in *Artificial Intelligence*, Vol. 173, pp. 789–816 (2009)
- [Shinyama 00] Shinyama, Y., Tokunaga, T., and Tanaka, H.: "Kairai" - Software Robots Understanding Natural Language, in *Proc. of 3rd International Workshop on Human Computer Conversation*, pp. 158–163 (2000)
- [Winograd 87] Winograd, T. and Flores, F.: *Understanding Computers and Cognition: A New Foundation for Design*, Addison-Wesley (1987)