

クイズ番組に挑戦したコンピュータの開発から学んだこと

統計数理研究所 公開講演会

A virtual game show set. On the left is a Jeopardy! board with six columns and five rows of dollar amounts. In the center, a host in a suit stands behind a podium. Behind the host is a large screen displaying a blue and green abstract graphic with a person's silhouette. The floor has a circular light projection.

JEOPARDY!	JEOPARDY!	JEOPARDY!	JEOPARDY!	JEOPARDY!	JEOPARDY!
\$200	\$200	\$200	\$200	\$200	\$200
\$400	\$400	\$400	\$400	\$400	\$400
\$600	\$600	\$600	\$600	\$600	\$600
\$800	\$800	\$800	\$800	\$800	\$800
\$1000	\$1000	\$1000	\$1000	\$1000	\$1000

日本アイ・ビー・エム株式会社
東京基礎研究所
武田浩一

2011年11月2日

目次:

1. IBM Researchが取り組むグランド・チャレンジ

2. Jeopardy!とWatsonについて

- 米国の人気クイズ番組 **Jeopardy!**
- 質問応答システム**Watson**

Watsonが実現した知的処理と解決した技術的課題

3. 質問応答の仕組みと技術的な意義

- 高度な言語処理
- **Watson**の処理の流れと、それを実現するハードウェア

4. Lessons Learned

2011年2月14-16日、IBMの研究部門が開発した質問応答システム Watsonが米国の人気クイズ番組「Jeopardy!」に挑戦しました

“Jeopardy!” とは

- 1964年から続いているアメリカのクイズ番組
- 月曜～金曜の夕方、既に9000回以上放映
- 歴史・科学・スポーツなど、幅広い分野から出題
 - 知識が問われる正統派のクイズ
 - 言葉遊び的なパズルも

- 解答者3人が獲得金額を競う形式

Watsonの対戦相手:

- Ken Jennings氏 (最多連勝記録保持者)
- Brad Rutter氏 (累積最高賞金記録保持者)



Jeopardy! の問題と解答の方法

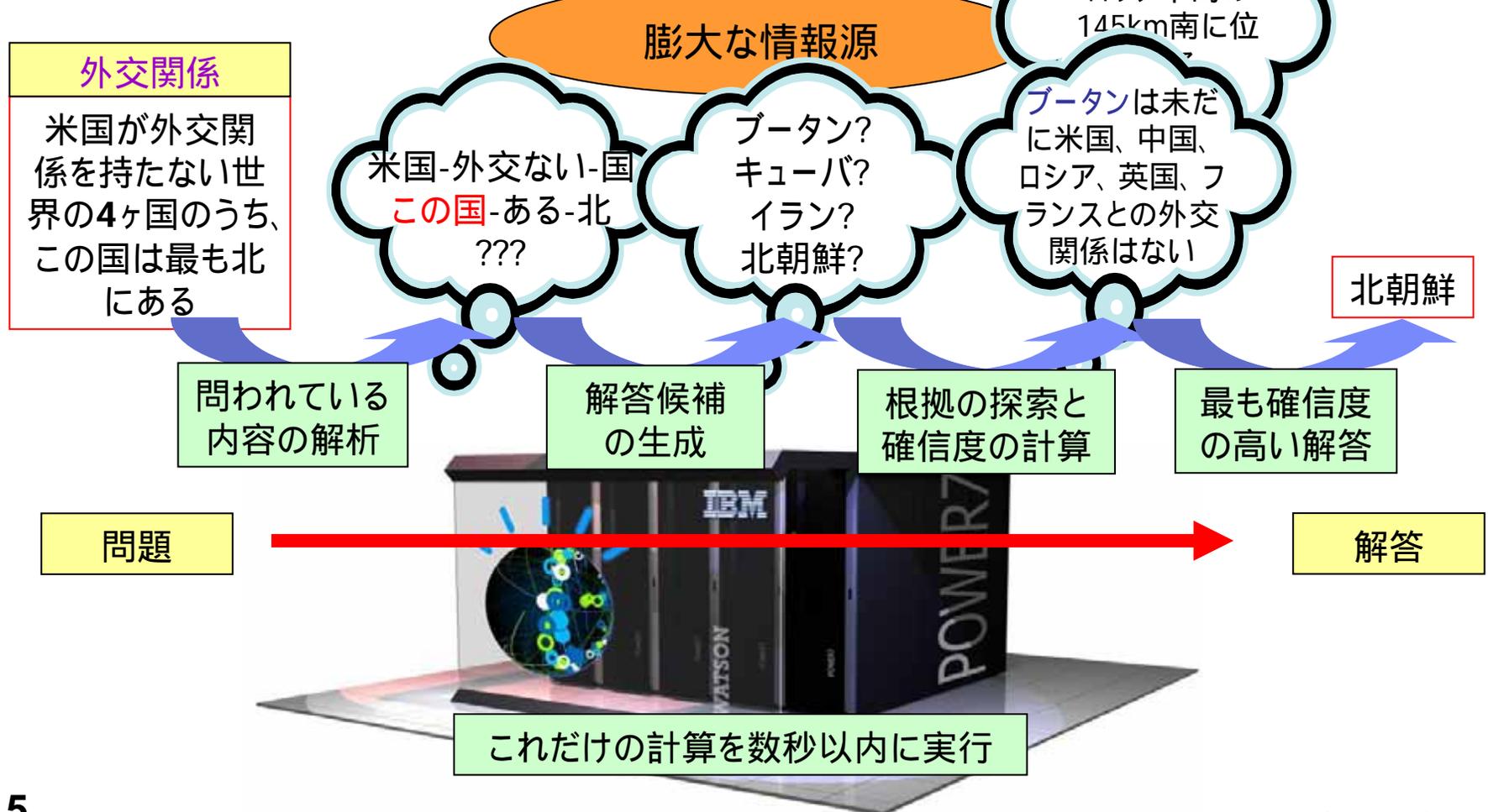
- 解答者の一人が問題を選択
 - 各問題に賞金が付与される
 - 例: 「外交関係」の \$1,000
- 司会者が問題を読み上げる
 - 同時に問題が表示される
- 解答者は早押しで解答
 - 司会者が読み終わった後に
- 解答が正しければ賞金を得る
- 不正解なら同じ金額を失い、他の解答者に解答権が移る



米国が外交関係を持たない世界の4ヶ国のうち、この国は最も北にある

質問応答システムWatsonとは？

問題(文)の内容を分析して、事前に収集された大量のテキスト情報から問題の解答候補とその根拠・確信度を計算し、高い確信度の候補が得られた場合に解答する、という一連の知的処理を高速に実行するコンピューター・システム



Watsonが実現した知的処理

- 問題をテキストで受け取り、その内容から、解答の候補を列挙する
- 蓄えられた文書と照らし合わせて各解答候補の確からしさを検証する
- 一定以上の確信を得た時にボタンを押す
- 音声で答えを読み上げる
- 賭け金やパネルの選択など、ゲームの進行上必要な判断をする
- **Watson**には実装されていない機能：
 - 聞くこと(音声認識)
 - 見ること(画像認識)
 - インターネットへの接続(対戦時)
 - 人間のような感情や直感
(人間を模倣するのではなく、問題の正しい解答を計算する機能に注力)



実際の問題の例と、Watsonの解答

DIALING FOR DIALECTS

WHILE MALTESE
BORROWS MANY
WORDS FROM ITALIAN,
IT DEVELOPED FROM A
DIALECT OF THIS
SEMITIC LANGUAGE

問題: マルタ語はイタリア語から多くの語彙を借りているが、それはこのセム語系言語の方言から発展した。



What is Arabic?

答え: アラビア語

問題文が非常に複雑

事前に問題を予測して

答えのデータベースを作るのは**不可能**
キーワード型の検索エンジンを用いても
答えを見つけるのは困難

そもそも検索エンジンが返すのは
文書集合であり、その中から答えを
探すことは非常に難しい

問題を解くために必要な言語処理

While Maltese

it developed from a dialect of this Semitic language

問われている内容の解析

答えるべきものは

“this Semitic language” (セム語族の言語)

代名詞の照応解析(指示対象の解釈)

“It” は “Maltese” (マルタ語) を指す

(Italian(イタリア語)だと解釈すると誤る)

手がかりの検出

(= 解答の候補)

- はセム語族の言語
 - マルタ語は、 の方言から発展した
- この「手がかり+候補」を満たす根拠を調べる

問題の前半部分は解答に直接関係しない

問題: マルタ語はイタリア語から多くの語彙を借りているが、それはこのセム語系言語の方言から発展した。

根拠付け：キーワードの一致に頼ると誤答につながる

問題文

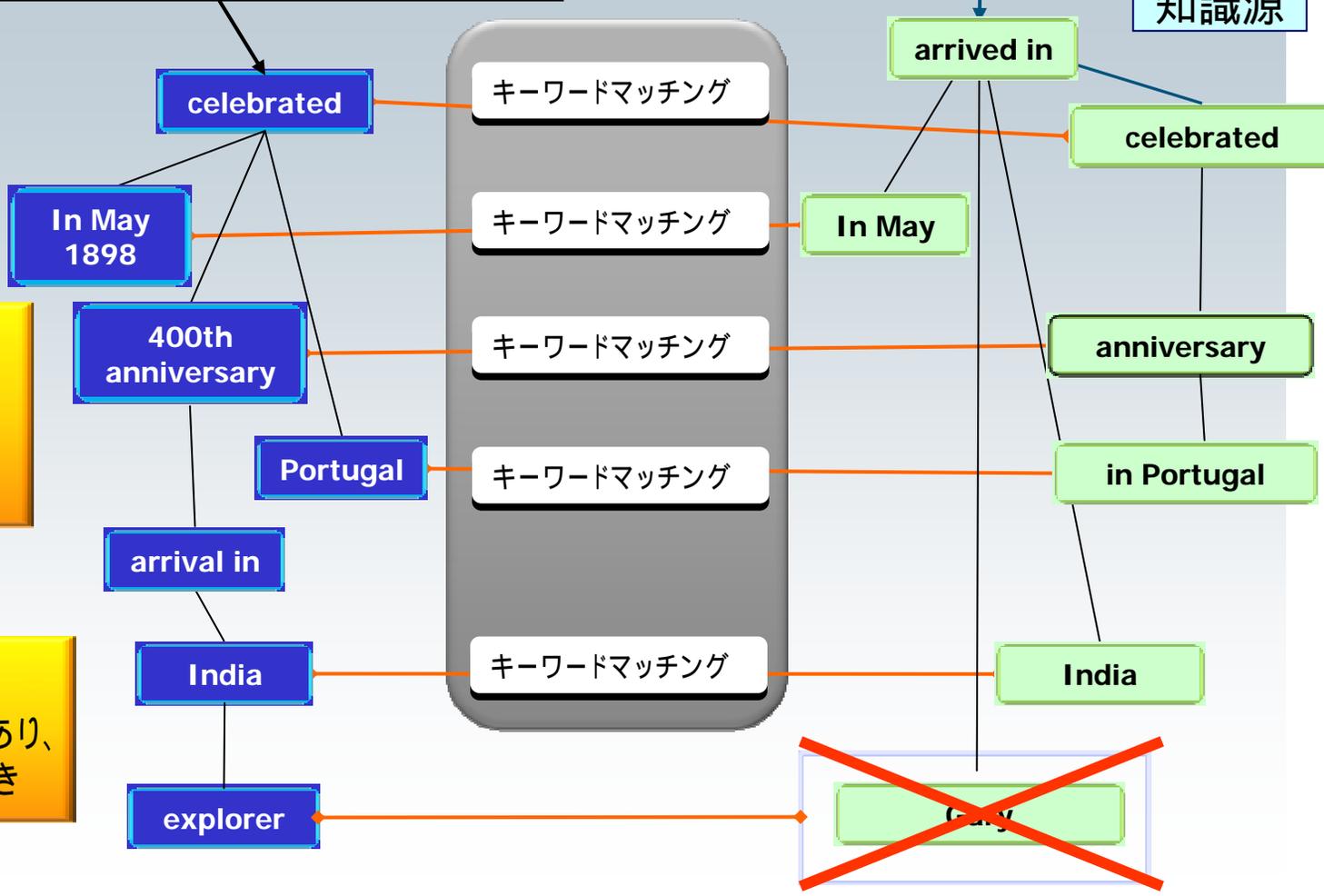
In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.

In May, Gary arrived in India after he celebrated his anniversary in Portugal.

知識源

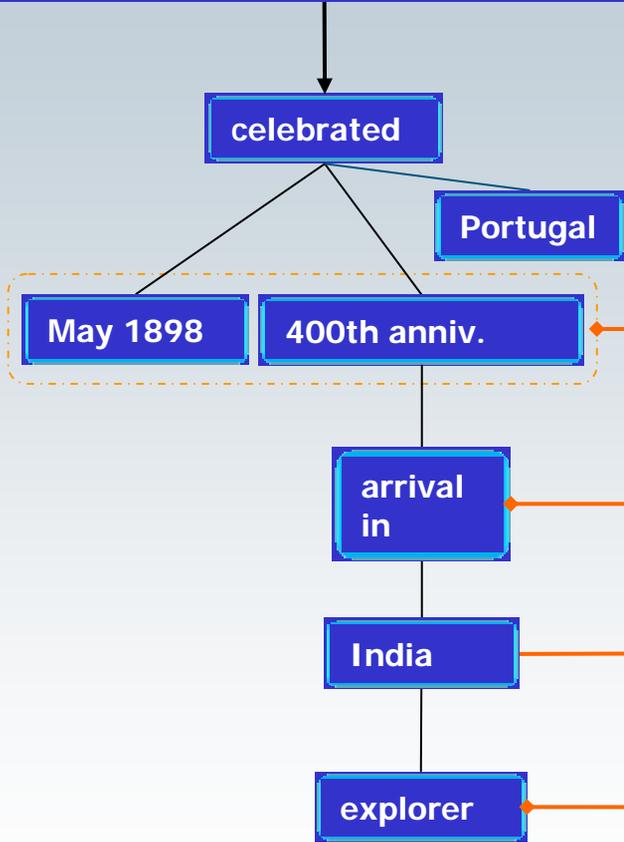
「Gary」が解答となることの「根拠」が大量に見つかってしまう

「キーワードの一致」はあくまでも弱い根拠であり、重み付けが低くなるべき

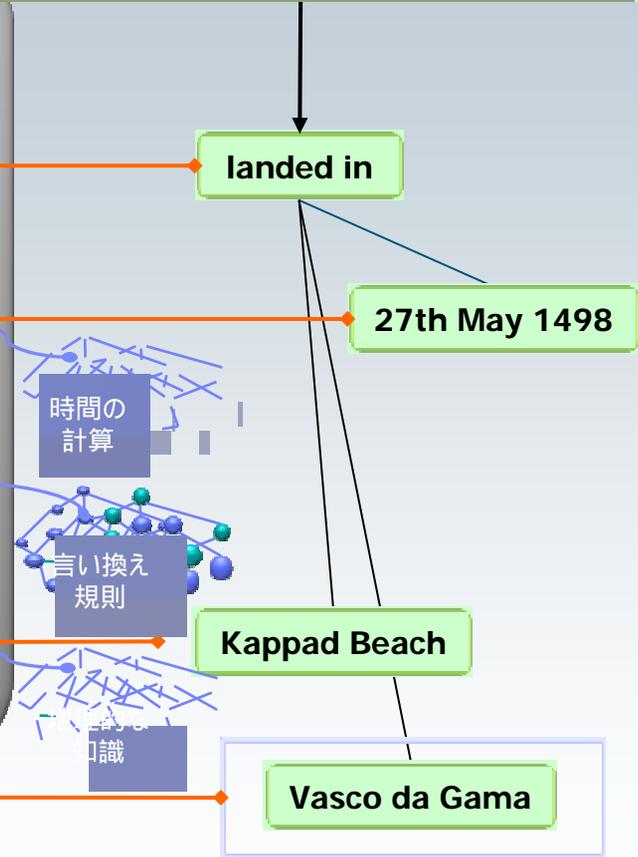


根拠付け：意味を考慮したマッチングにより正確な答えを得る

In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.



On the 27th of May 1498, Vasco da Gama landed in Kappad Beach



広範囲を深く探索
 たくさんの仮説を吟味
 根拠を集めて判断
 多くの推論アルゴリズム

同義計算
 統計的な換言処理
 地理的な類推

時間の計算
 言い換え規則
 常識

• 意味的な一致は「強い根拠」になるが、発見・数値化が難しい
 • どんな根拠も100%信頼できるものではない

Watsonを支える技術は今後も多くの応用が期待されます

より迅速で正確な医療診断支援のためのシステム

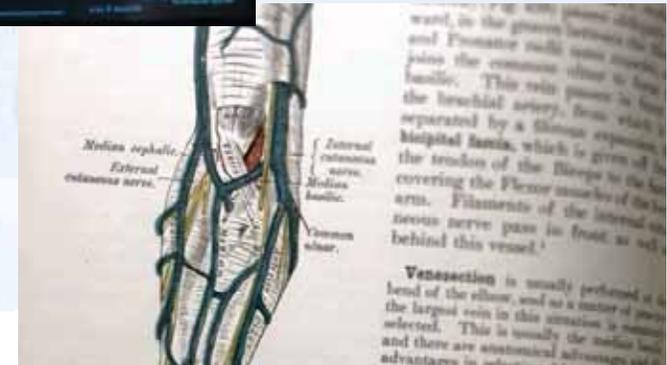
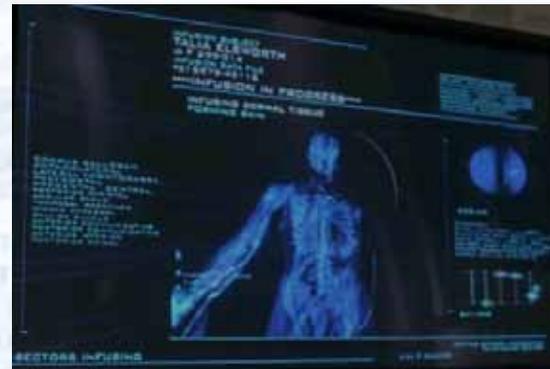
潜在的な薬物間相互作用の検査のためのシステム

弁護士や裁判官による過去の判例の参照のためのシステム

金融分野の仮説シナリオと法令順守のためのシステム

Major Currencies - Real-Time

Currency	Bid	Ask	Change	High	Low	Time
EUR/USD	1.2183	1.2186	-0.0018	1.2359	1.2178	19:31:39
USD/JPY	89.962	89.975	-0.485	90.67	89.81	19:31:29
GBP/USD	1.4407	1.4409	-0.0001	1.4443	1.433	19:31:37
USD/CHF	1.15	1.16029	0.0029	1.1624	1.1511	19:31:34
USD/CAD	1.0665	1.06688	-0.0011	1.0744	1.0578	19:31:25
AUD/USD	0.8239	0.82389	-0.0001	0.8389	0.8198	19:31:38
EUR/JPY	109.39	109.435	-0.11	111.81	109.39	19:31:31
EUR/CHF	1.4133	1.4138	-0.0113	1.4293	1.4129	19:31:36
GBP/JPY	129.32	129.47	0.63	130.55	129.04	19:31:27
GBP/CHF	1.6712	1.6716	0.0037	1.6745	1.6582	19:31:34
CHF/JPY	77.447	77.49088	-0.1088	78.471	77.37	19:31:40
NZD/USD	0.8919	0.8929	0.0012	0.8743	0.8601	19:31:35
USD/ZAR	7.7351	7.78005	0.0005			



エビデンスに基づく診断の支援

症状



家族の既往症
両親の既往症
薬物治療
検査/発見



所見/仮説



広い範囲のエビデンスを考慮して診断を支援し、医療の質の向上とコスト低減に貢献



大規模な医療関連情報(文書、論文、記事、データベースなど)



Lessons Learned

1. 困難な目標に果敢に挑戦することで得たものは多かった
2. **NLP**ツールを積極活用したシステム開発手法が有効であった
3. いま利用可能なデジタル情報の質と量の凄さを実感した
4. 別の分野に適用するには、今回と同程度の質・量の情報が必要

参考情報

1. Watson紹介Webページ(英語)

<http://www.ibmwatson.com>

他に個別の技術についての論文発表が数件あります

https://researcher.ibm.com/researcher/view_page.php?id=2107

2. Watson紹介Webページ(日本語)

<http://www.ibm.com/ibm/jp/lead/ideasfromibm/watson/>

3. “Building Watson: An Overview of the DeepQA Project” AI Magazine, 2010, (技術解説論文)

<http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/2303>

4. 「IBM奇跡の“ワトソン”プロジェクト - 人工知能はクイズ王の夢を見る」 “Final Jeopardy”, スティーヴン・ベイカー (著), 土屋 政雄 (訳) 出版社: 早川書房 (2011/8/25), ISBN: 978-4-15-209236-6

質問応答システム“ワトソン”がクイズ番組に挑戦！

クイズ王との競演で“ワトソン”が魅せた
コンピューターの大なる可能性

世界が注目、自然言語による
質問応答システムは新たな水準へ

歴史的な1日となった2011年2月16日
2011年2月16日(米国時間)、コンピューターの歴史に新たな1
ページが加わりました。IBMワトソンの4年間にわたる研究成
果である質問応答(QA)システム「Watson」(以下「ワトソン」)が、
米国の人気クイズ番組「Jeopardy!」(以下「ジョパディ!」)に挑
戦し、2ゲームを通じて、ワトソンが最高金額を獲得しました。
当時賞金100万ドルの全米放送番組とされます。ワトソ
ンは研究者の集約の結果でもあり、「ジョパディ!」の最新対人型である、と出場者たちは今
回の対戦を誇りにしています。

ワトソンとは
ワトソンは、質問応答技術の更なる向上を目的に、自然言語処理技術を中心に強化することを
目的に設計されました。「ジョパディ!」で出題されるバラエティーに富んだ複雑な問題に対して、
100万問の単語に相当する自然言語で書かれた情報の断片を分析し、短時間で最も適し
た解答を提出する分析エン지니어リングシステムです。この研究開発プロジェクトには、IBM美
国基礎研究用から60名の研究者が参加し、グローバルの技術力を結集して取り組んでいま
した。