



オフショアソフトウェア開発における異言語 文書理解支援システムの実験的評価

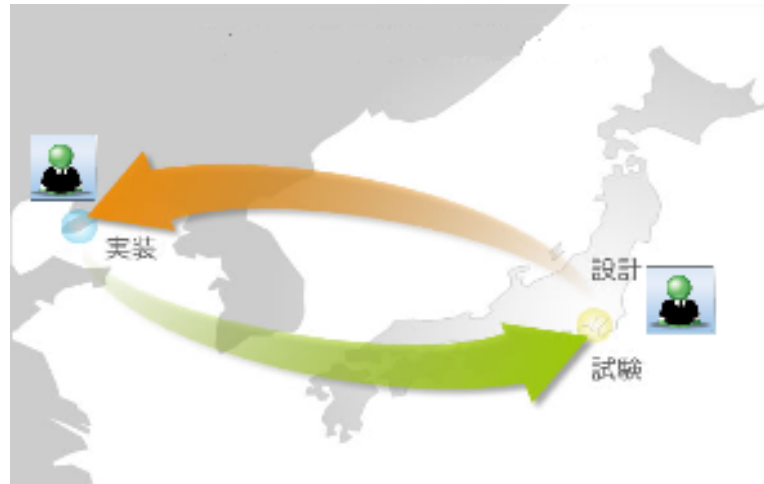
神戸大学大学院システム情報学研究科
程 祥瑞, 松本 真佑, 中村 匡秀

2011年11月24日
サービスコンピューティング時限研究専門委員会 第5回研究会

オフショア開発

- ❁ システム開発プロセスの一部を海外の企業に委託する開発形態
 - ⦿ コストダウン
 - ⦿ 開発資源(人材, ノウハウ)の外部調達
- ❁ 実例: 米国のある銀行
 - ⦿ 新システムの導入に国内で\$1,000万かかると見込まれた
 - ⦿ 中国にあるソフトウェア会社に委託し, 10ヵ月, \$9万で開発に成功した

オフショア側
(中国)
・実装



オンショア側
(日本)
・設計
・試験

オフショア開発における問題点

- ❁ コミュニケーション問題

- ⦿ 言語の違いから発生する問題
- ⦿ 例) 英語がわかるけど, 英文資料を理解すると時間がかかる

- ❁ 文化差異

- ⦿ 文化や習慣の違いから発生する問題
- ⦿ 例) 時間に対する感覚

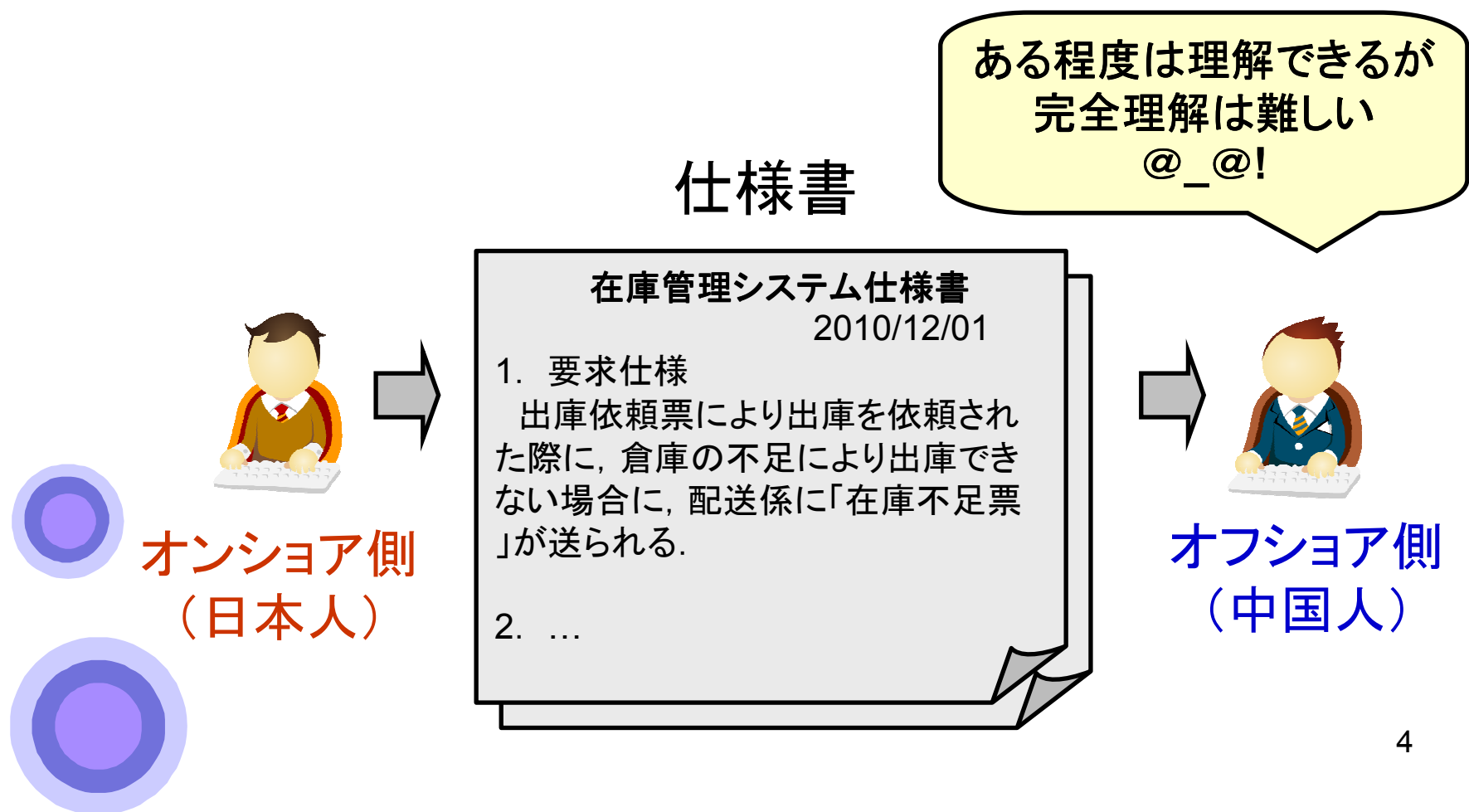
- ❁ 本研究で解決するのは

「コミュニケーション問題」



コミュニケーション問題1

❁ P1. オフショア側の文書理解問題



コミュニケーション問題2

❁ P2. オンショア側の理解度把握問題

どの程度理解してるのか
よく分からない
@_@!

仕様書



オンショア側
(日本人)

在庫管理システム仕様書
2010/12/01

1. 要求仕様
出庫依頼票により出庫を依頼された際に、倉庫の不足により出庫できない場合に、配送係に「在庫不足票」が送られる。
2. ...



オフショア側
(中国人)

研究目的とアプローチ

❁ 研究目的

- ❁ 2つのコミュニケーション問題を解決し、オフショア開発を円滑に進めるように支援すること

❁ アプローチ

- ❁ 異言語間開発の為にソフトウェアドキュメント共有システムを開発する

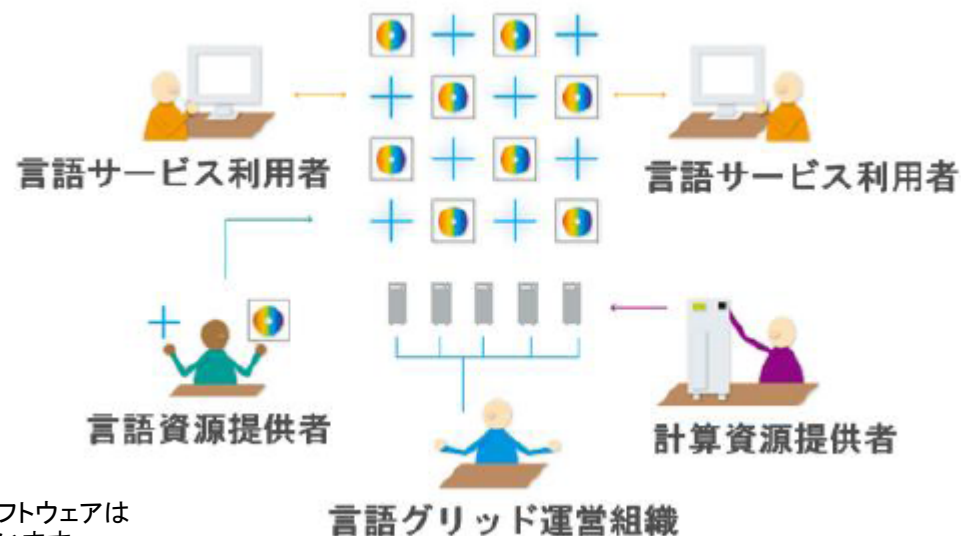
1. 言語グリッドウェブサービスを用いた機械翻訳
2. ドキュメント理解度のフィードバック・共有



言語グリッド ※



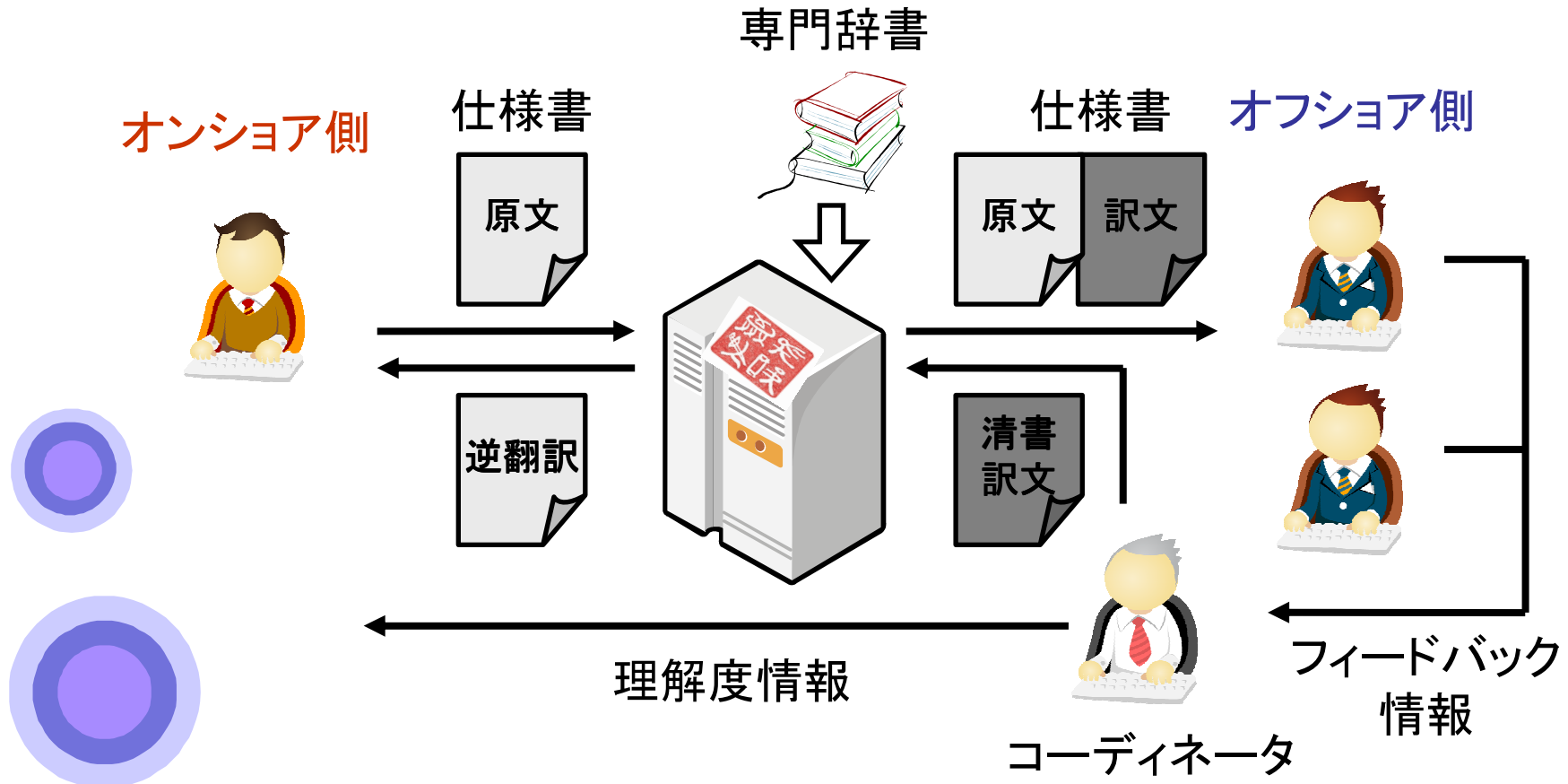
- ❁ 辞書や機械翻訳などの言語資源を，インターネット上から利用できる多言語サービス基盤
- ❁ ユーザは様々な言語サービスを組み合わせる利用することができる
 - ⦿ 日中翻訳サービスと中チベット翻訳サービスを組み合わせることで，日チベット翻訳を実現



※言語グリッドを実現する基本ソフトウェアは
NICTで研究開発され公開されています

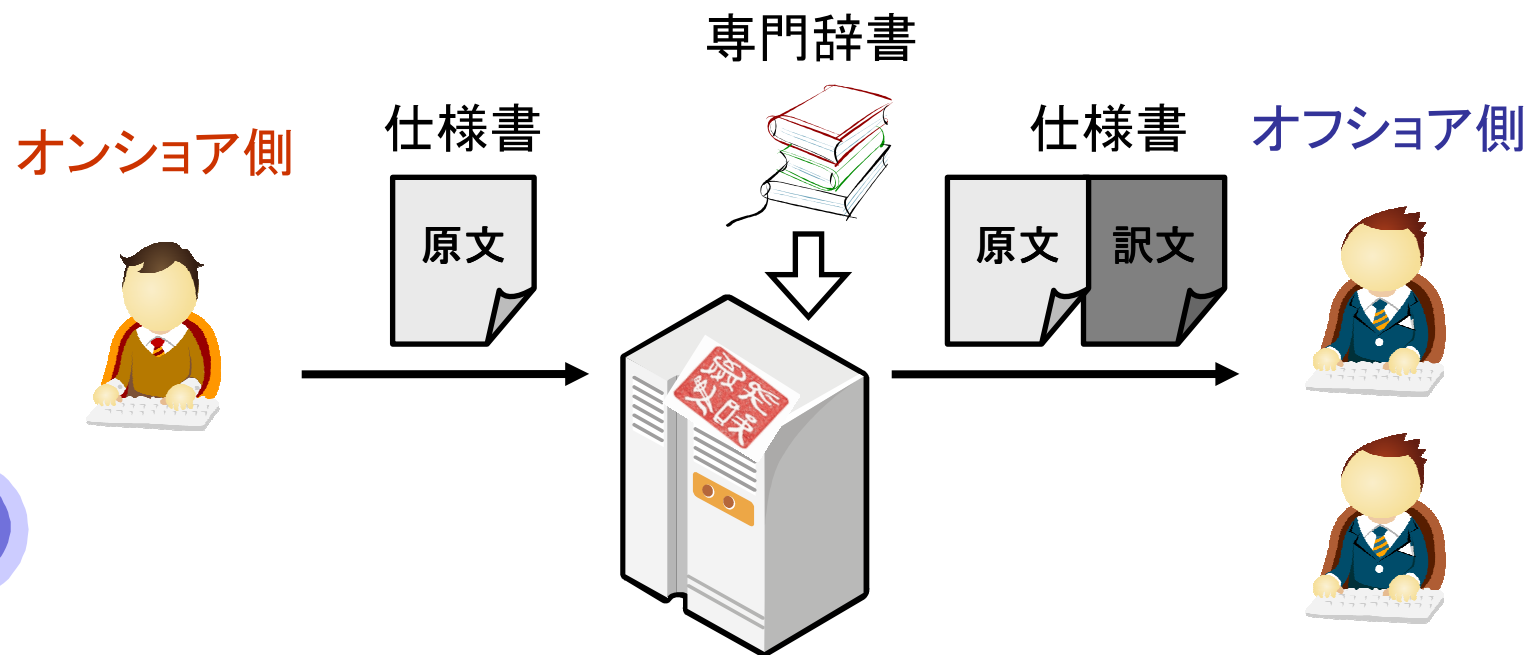
提案手法

- ❁ 仕様書の翻訳プロセス
- ❁ オフショアからの理解度フィードバック



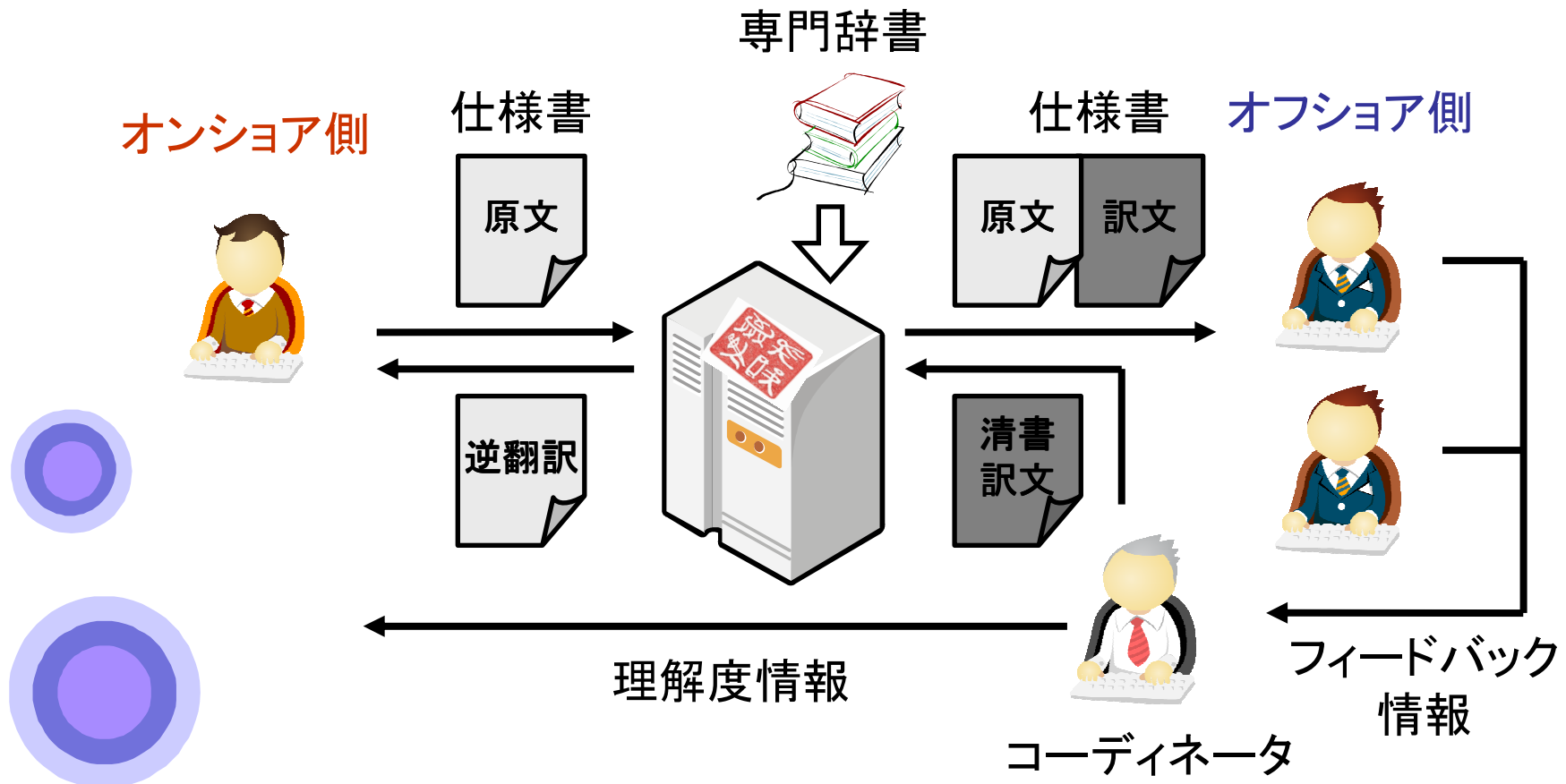
仕様書の翻訳プロセス

- ❁ 仕様書の原文と訳文をオフショア側に提供する。
 - ▣ P1. オフショア側の文書理解問題に対処.

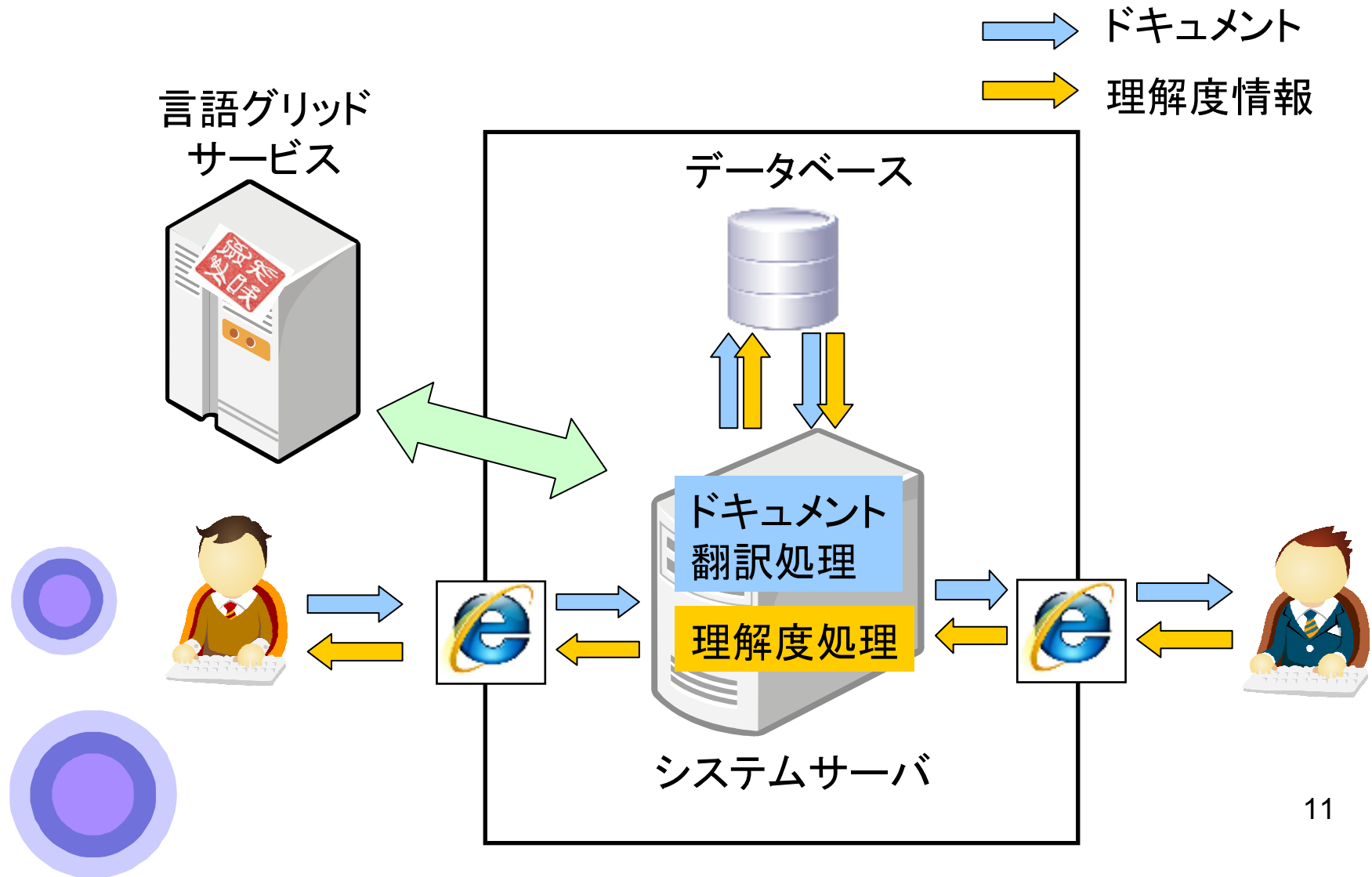


オフショアからの理解度フィードバック

- ❁ オフショア側の理解度情報をフィードバックする。
 - 📌 P2. オンショア側の理解度把握問題に対処



提案システムのアーキテクチャ



プロトタイプシステム

- ❁ DBサーバー: MySQL
 - ❁ サーバー側: APACHE, PHP
 - ❁ クライアント側: HTML, JavaScript
 - ❁ プラットフォーム: XOOPS, 言語グリッドToolbox
-
- ❁ 完成した部分: ドキュメント翻訳, ドキュメント閲覧
 - ❁ 未完成機能: 専門辞書, 理解度フィードバック



システムのトップ画面

The screenshot shows the top page of the Language Grid Toolbox web application. The browser window title is "Multilingual Document Understanding Supporting System". The page header includes the text "Multilingual Document Understanding Supporting System" and a navigation menu with icons for "Top", "Communication", "Translation", "Creation", "Offshore", "Onshore", and "Setting". The "Top" button is highlighted in green. The main content area features a world map background and a text box describing the software. The text reads: "Language Grid Toolbox is an open-source software, developed by the National Institute of Information and Communication Technology. The software supports multilingual communication using the Language Grid." Below this, it lists features: "A wide variety of communication support functions" and "Functions' expandability". At the bottom, there is an "Information" section with release dates for versions 2.3 and 2.2.

Welcome, admin | [Logout](#)

Multilingual Document Understanding
Supporting System

Top Communication Translation Creation Offshore Onshore Setting

Language Grid Toolbox is an open-source software, developed by the National Institute of Information and Communication Technology. The software supports multilingual communication using the Language Grid.

[Features of Language Grid Toolbox]

- **A wide variety of communication support functions** The software provides a variety of functions including Text Translation, Multilingual Web Creation, and Language Resource (Multilingual Dictionary, Parallel Texts, Q&As, and Glossaries) Creation. Combining machine translators and the created dictionaries improves the translation quality.
- **Functions' expandability** To meet various needs of multilingual communities, Toolbox allows addition of new functions. The software is designed to allow development of new functions by combining existing functions, and it lowers the development cost.

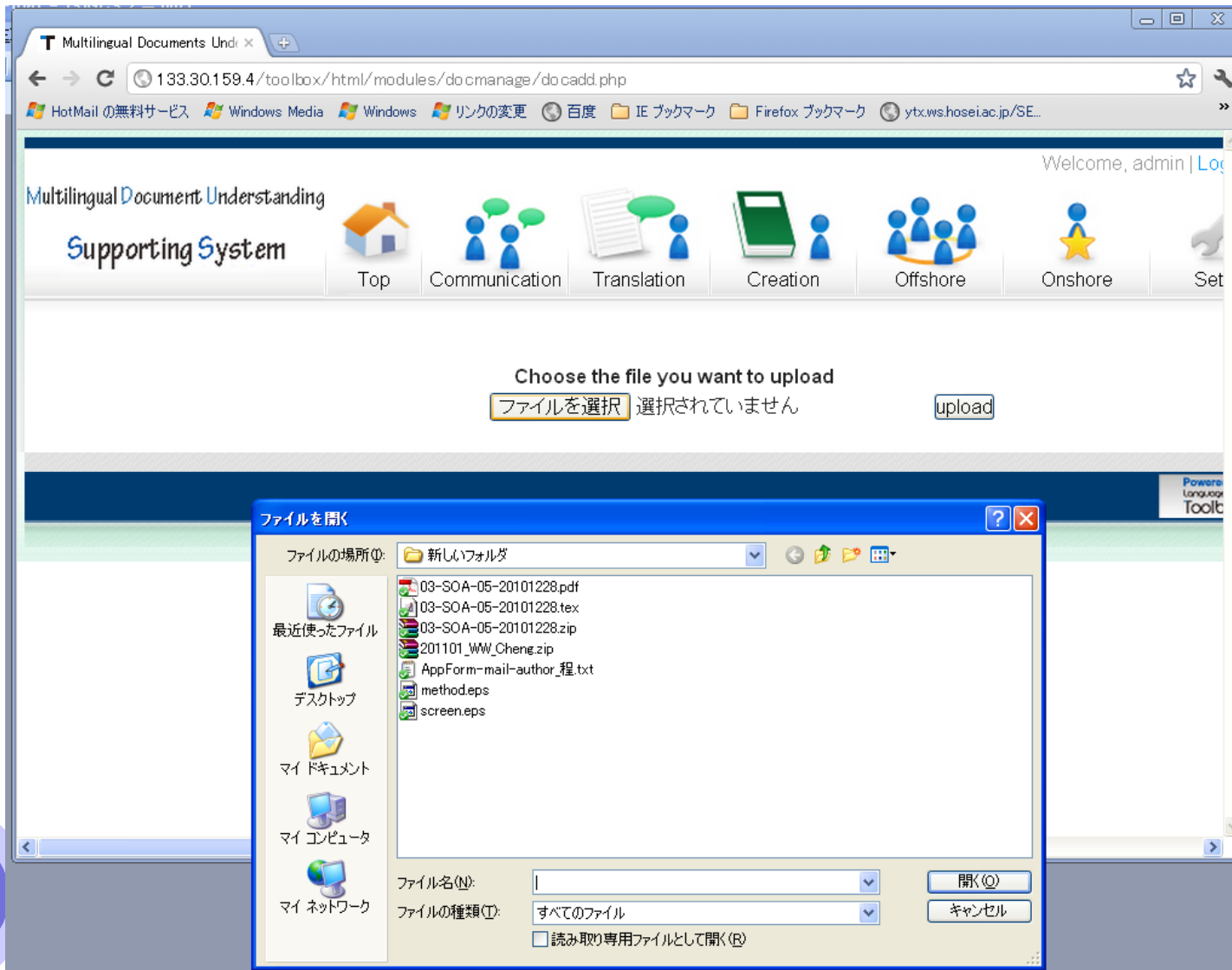
Please visit the [Toolbox Support Site](#) for the information of free hosting service and downloading the source code.

* Language Grid Toolbox is tested on Internet Explorer 7 and 9, and Mozilla Firefox 3.6

Information

2010.4.1 Language Grid Toolbox version 2.3 has been released.
2010.2.26 Language Grid Toolbox version 2.2 has been released.

ドキュメントの追加画面



システムのドキュメント閲覧画面

ID	Orinal Text	Translated Text	Understanding
1	酒管理プログラム要求仕様書	酒管理程序要求说明书	5 ▼
2	1. 要求仕様	1. 要求配置	5 ▼
3	ある酒屋では、すべての在庫(酒)を幾つかのコンテナに載せて、一つの倉庫に保管している。	在某酒店,为几个的集装箱登载全部的库存(酒),为一个仓库保管着。	5 ▼
4	「受付係」の仕事は、倉庫で在庫の積み下ろしを行う「倉庫係」、及び、客の注文を聞いて客に酒を届ける「配送係」との間で数種類の伝票をやりとりしながら、倉庫内の在庫を管理することにある。	「接待员」的工作,在在仓库进行库存的装载卸的「仓库员」,和,听客人的订购送交酒给客人的「发送员」间一边交换数种类传票,一边在于管理仓库内的库存。	5 ▼
	1.1 受付係への入力	向1.1接待员的输入	

システムの理解度確認画面

ITEMID	TEXT	UNDERSTANDING	DATE
19	配送係への出力である在庫不足票と、倉庫係への出力である出庫指示票の区別は行わず、標準出力へ出力するものとする。	2	2011-09-25
20	1.2 受付係の出力	5	2011-09-25
21	受付係の出力は全て、標準出力へなされるものとする。	5	2011-09-25
22	(I2) 出庫依頼票	5	2011-09-25
23	客から酒の注文があると、配送係から「出庫依頼票」が送られてくる。	4	2011-09-25
24	出庫依頼票には、出庫すべき酒の銘柄(1 銘柄のみ)、数量、及び、送り先名が記入されている。	4	2011-09-25
25	出庫依頼票 送り先 名銘柄 数量	5	2011-09-25
26	積荷票 コンテナ 番号 搬入年月 搬入日時 銘柄数(銘柄, 数量) の繰り返し	5	2011-09-25
27	さらに、積載されている銘柄と銘柄毎の数量(積載量) が積載銘柄数だけ繰り返し記入されている。	5	2011-09-25
28	1 台のコンテナに酒を10 銘柄まで混載できるため、積荷票には、搬入されたコンテナの番号、搬入年月、搬入日時と共に、	3	2011-09-25
29	(I1) 積荷票	5	2011-09-25
30	酒を積んだコンテナが倉庫に搬入されると、倉庫係から「積荷票」が送られてくる。	5	2011-09-25

翻訳精度に関する予備実験

* オリジナル仕様

- ☐ 「顧客から酒の注文がある場合、配送係は在庫依頼票を発行して、受付係に送ります。」

* 翻訳した仕様

- ☐ 「如果从顾客有酒的订购，发送员发行发出委托票，送接待员。」

* 逆翻訳した仕様

- ☐ 「もし顧客から酒の予約購入があるならば、発送(送信)員発行は出して切符(チケット)を頼んで、接待員に送る。」

訳文は読みにくいけど、意味が分かる

実験の説明

❁ 被験者

- ❏ 中国人8名
- ❏ 日本語ビジネスレベル4名, 日常会話レベル4名
- ❏ IT会社社員4名, 学生4名

❁ 実験内容

- ❏ 日本語の酒屋在庫管理システムの仕様書を読んでもらい, その理解状況を確認する.
- ❏ 提案システムを使う場合(システム)と, 使わない場合(直接)の2通り実施



実験の説明(続)

* 評価方法

- ▣ 仕様書に関する問題に対する回答で採点
(計6問, 各問5段階評価)

- ▣ アンケート

* 2回目の学習効果に関する対応

- ▣ 2回目の実験は, 一週間以上開ける
- ▣ 半分の被験者 直接→システム
- ▣ 半分の被験者 システム→直接

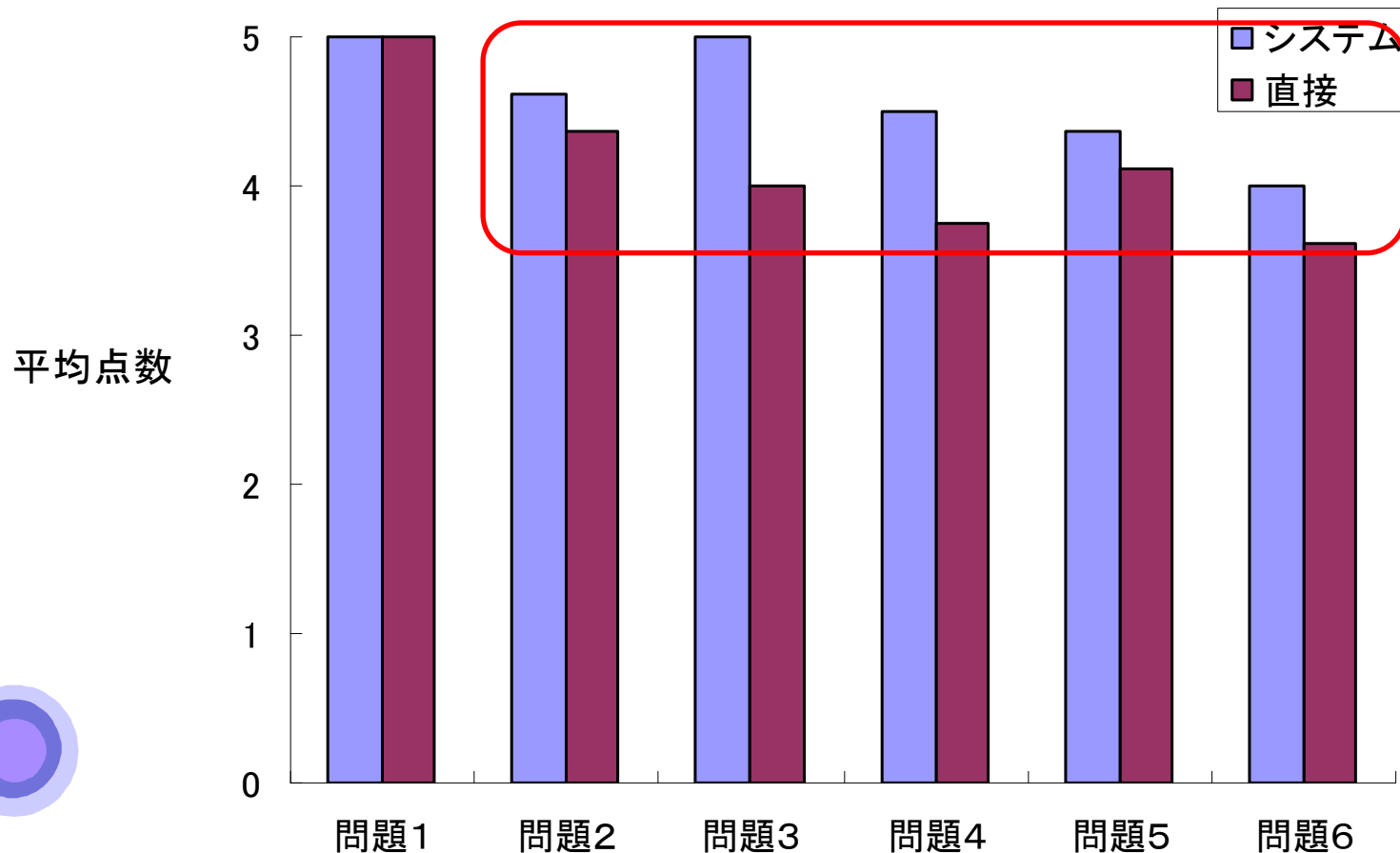


仕様書理解の確認問題

- Q1. 1つのコンテナに、最大で何種類の銘柄が入れられますか？
- Q2. 在庫不足票はどの係からどの係に送られますか？
- Q3. 2011年8月20日20時にコンテナが搬入された場合、搬入年月と搬入日時はどのようになりますか？
- Q4. 倉庫に15個のコンテナがあります。各コンテナの中に100本SUN焼酎があります。1200本の焼酎の出庫依頼が来た時、受付係からどの伝票が何枚発行されますか？（正常出庫）
- Q5. 倉庫に1000本のSUN焼酎があります。順番に焼酎800本、300本、100本の3回の出庫依頼を受けました。それぞれの出庫依頼に対して、受付係からどんな伝票が発行されますか？（在庫不足）
- Q6. 下記のプログラムの表示は、ある出庫依頼がきたときの出力結果です。この出力結果から何回の出庫が行われますか？また各出庫に対してそれぞれ何本の酒が出庫されますか？

※Q1が一番簡単、徐々に難しくなり、Q6が一番難しい

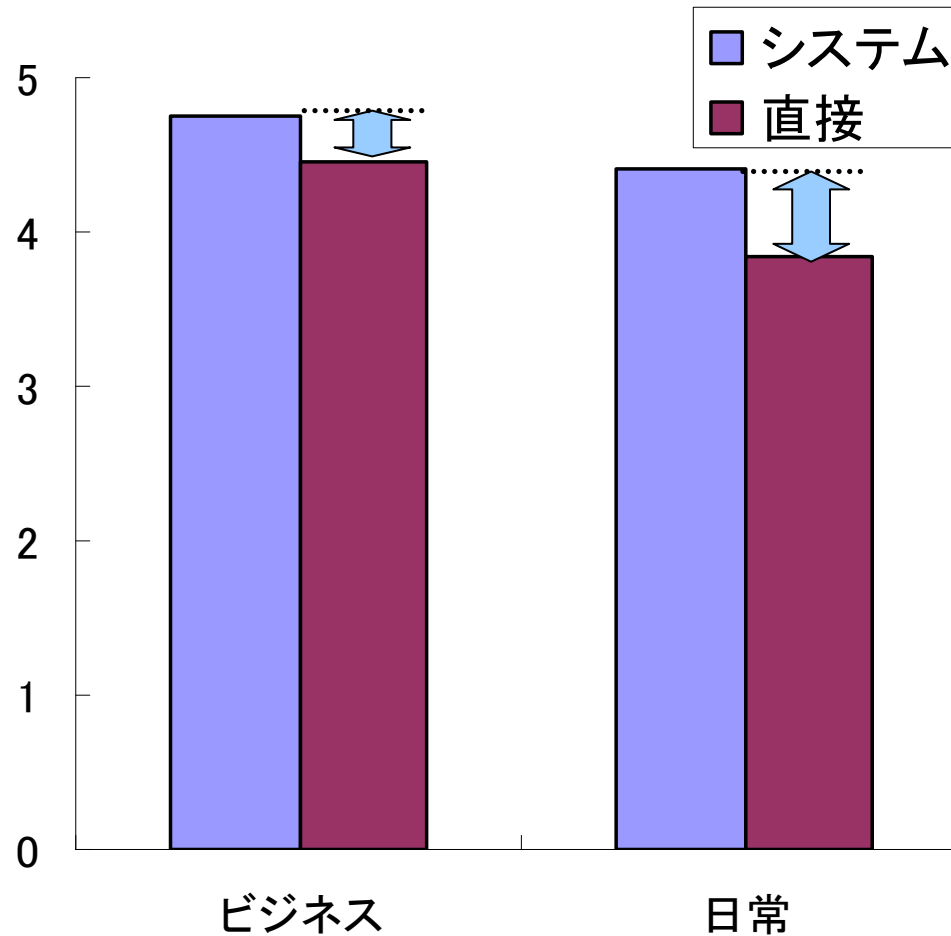
実験結果(問題別採点結果の被験者平均)



簡単な問題を除き, 提案システムによって
概ね高い理解度が得られた
→ 提案システムが仕様書理解に役立つ見込みを得た

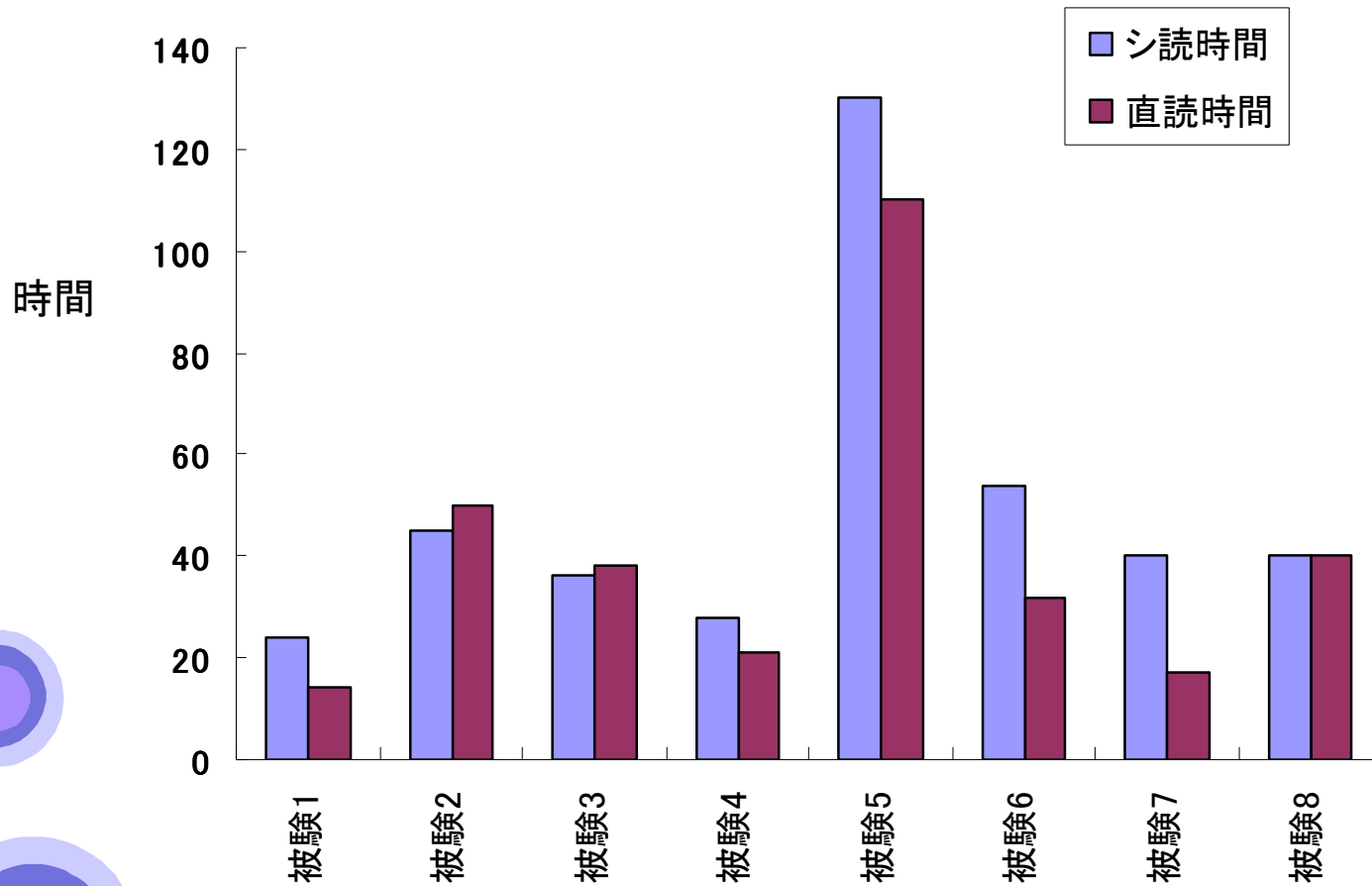
実験結果(日本語レベル別の採点結果)

総平均点



特に日本語レベル低い人の理解幫助に効果がある

実験結果(理解に要する時間)



システムを使用する場合は、仕様書を直接読むより、時間がかかる

アンケート結果

❁ 良かった所

- ⊙ 機械翻訳であっても母国語に変換されていると安心.
- ⊙ 翻訳があるので, 難しい単語と複雑な表現が理解し易い.
- ⊙ システムを使うことで, 文書の全体像が把握できる.
- ⊙ 項目ごとに記録した理解度は, 自分の読み直しとPMへのフィードバックに役立ちそう.

❁ 悪かった所

- ⊙ 翻訳精度によっては, 誤解も発生しやすい
- ⊙ 長文の翻訳精度が低い

❁ その他

- ⊙ 検索機能とハイライト機能がほしい
- ⊙ 翻訳インターフェイスの改善要求

まとめ

- ❁ オフショア開発における異言語文書の理解支援システムを試作した.
- ❁ 被験者実験により, 有効性を確認した.
 - ▣ 提案システムを用いた方が理解度が高かった
 - ▣ 母国語の安心感, 難語・長文の理解, 全体の把握
- ❁ 今後の課題
 - ▣ システムのインターフェイスの改善
 - ▣ オンショア側の支援機能の実装
 - ❁ 専門辞書
 - ❁ 理解度フィードバック機能