

質問応答:ファクトイドから実世界・
大学入試歴史科目問題まで

横浜国立大学

澁木 英潔

2016/11/10

インタラクティブ情報アクセスと可視化マイニング研究会

目次

質問応答システムの歴史—なぜ大学入試歴史科目問題を扱うのか？—

大学入試問題への取り組み

語句記述問題(factoid型質問)

論述問題(non-factoid型質問)

NTCIR QA Lab Taskへの招待

質問応答の歴史

質問応答とは

- ・ 自然言語で表現された質問に対して適切な回答を与える

自然言語理解や対話システムなどで黎明期から研究

- SHRDLU (1972) ⇒ クローズドな世界
- FAQ Finder (1997)

オープンメイン質問応答が主流

- TREC-8 QA Track (1999)
- NTCIR-3 QAC Task (2001)
- IBM Watson (2011)



実世界を考慮して

NTCIR-5 CLQA Task (2004)

- 。多言語質問応答

TREC-2015 Live QA Track

- Yahoo Answersの質問に実時間(1分以内)で回答

NTCIR-11 QA Lab Task (2013)

- ・大学入試歴史科目(世界史)問題に回答

ロボットは東大に入れるか
(東ロボ)プロジェクトとも連携

⇒ なぜ(クローズドな)大学入試歴史科目問題を扱うのか？

NTCIR QA Lab task

- 情報検索、文書要約、情報抽出、質問応答などの情報アクセス技術を
促進するための評価ワークショップ

- Organizers**
- 神門典子(国立情報学研究所)
 - 石下円香(国立情報学研究所)
 - 三田村照子(カーネギーメロン大学)
 - **洪水英潔**(横浜国立大学)
 - 阪本浩太郎(横浜国立大学)

Advisers

- 森辰則(横浜国立大学)
- 狩野芳伸(静岡大学)

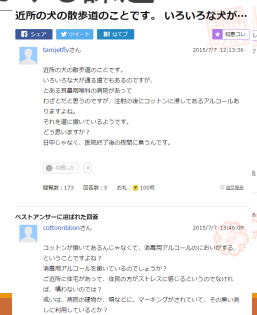
質問応答における課題

知識源の不備

- ⇒ オープンドメイン

質問内容の理解

適切な回答の生成



入試問題の例

2009年度 東京大学 世界史 第3問 問(2)

フランス革命期、ジャコバン派の独裁体制が打倒され、穏和派の総裁政府が樹立されると、革命の徹底化と私有財産制の廃止を要求する一部の人々は、秘密結社を組織して武装蜂起を計画したが、失敗し弾圧された。**この組織**の指導者の名を記しなさい。

パンテオン・クラブ

A バブーフ

大学入試歴史科目問題を扱った取り組み

NTCIR-9 RITE (Recognizing Inference in Text) Task (2010)

- センター試験(世界史)の正誤問題(yes/no型質問)

ロボットは東大に入れるか(東ロボ)プロジェクト (2011)

- 初期はセンター試験(日本史、世界史)
- 2016成果報告会 日時:2016年11月14日(月)13:00-17:00 場所:一橋講堂

NTCIR-11 QA Lab Task (2013)

- ・センター試験(世界史B)と二次試験(論述問題含む)

入試問題を対象とする利点

- (1) 未解決課題を多く含み、その解決が技術の発展を促す
周辺文脈の理解や推論など比較的高度な知的活動
・ 論述問題における比較の厳密な解答の適切性
- (2) 模範解答が存在するのでシステムの評価がしやすい
- (3) 参考書の解説などに解答までの道筋が示されているのでアルゴリズムの参考にできる

NTCIR QA Lab taskの展開

NTCIR-11 QA Lab

- センター試験(世界史B)
- 二次試験
 - 東京大学、京都大学、北海道大学、早稲田大学、中央大学

NTCIR-12 QA Lab-2

- ロボットは東大に入れるか(東ロボ)プロジェクトと連携して、予備校主催の模試に参加

NTCIR-13 QA Lab-3

- 論述問題に焦点
 - <http://research.nii.ac.jp/qalab/> Email: qalab-admin@nii.ac.jp

ロボットは東大に入れるか2015 結果概況
進研模試 総合学力マーク模試・6月

ペネッセコーポレーション 2015年度進研模試 総合学力マーク模試・6月
(受験者総数 44.0万人)

	国語	数学		英語	理科	地理	8教科 5科目
		数学IA ¹⁾	数学IIB ²⁾	英語 (文法) ³⁾	物理 ⁴⁾	日本史 ⁵⁾	世界史 ⁶⁾
配点	200	100	100	200	50	100	100
算点比率	90	75	77	80	16	42	55
算点比率 学生 平均点	105.4	65.5	42.8	86.0	24.6	49.4	46.6
算点比率 偏差値	45.1	44.0	65.8	48.4	40.5	46.5	54.8
							66.5
							57.8
							416.4

*1 数学については、問題文を機械が理解可能な形式表現に変換する過程で、現在開発中および今後開発予定の部分（数式の意味解釈、文間の関係の解析など）に限り、一部、人手による追加・修正を加えた。

*2 物理では、人手で問題文を機械処理可能な形式表現へと変換した。

*3 5教科8科目文系型(国、数2科目、英筆記及びリスニング、地歴2科目、理1科目)での受験者11.6万人で集計した偏差値

(<http://21robot.org/>より転載)

ロボットは東大に入れるか2015 結果概況 第1回東大入試実戦模試										
聯合予備学校 2015/2016 第1回東大入試実戦模試(数学)										
(受験者数: 2,141人、合格者: 121人、合格者比率: 5.6%)										
【文系】 ¹⁾	問1 正答率(%)	問2 正答率(%)	問3 正答率(%)	問4 正答率(%)	問5 正答率(%)	問6 正答率(%)	問7 正答率(%)	問8 正答率(%)	問9 正答率(%)	合計
学生	20	20	20	20	20	80				
平均点 ²⁾	5.3	8.6	2.0	9.2	25.0					
偏差値 ³⁾	4	-	5	3	22					
偏差値 ⁴⁾	47.5	-	57.7	40.6	41.4					
偏差値 ⁵⁾	20	-	13	6	39					
偏差値 ⁶⁾	78.3	-	78.2	45.2	59.2					
【理系】 ¹⁾	問1 正答率(%)	問2 正答率(%)	問3 正答率(%)	問4 正答率(%)	問5 正答率(%)	問6 正答率(%)	問7 正答率(%)	問8 正答率(%)	問9 正答率(%)	合計
学生	20	20	20	20	20	20	120			
平均点 ²⁾	10.7	4.1	3.9	2.9	4.2	5.3	31.1			
偏差値 ³⁾	-	-	2	-	0	-	2			
偏差値 ⁴⁾	-	-	47.1	-	37.6	-	35.1			
偏差値 ⁵⁾	-	-	20	-	0	-	20			
偏差値 ⁶⁾	-	-	74.8	-	37.6	-	44.3			

¹⁾ 数字については、問題文を機械が理解可能な形式言語に変換する過程で、現在問題中および今後模試予定の部分(数式の意味解釈、文脈の理解の解析など)に限り、一部、人手による追加・修正を加えた。

²⁾ 算段、四則演算の平均点

³⁾ 途中過程に対する加点を行わず、解答結果の正否のみについての加点を行った場合の得点と偏差値

⁴⁾ 途中過程での得点を満点とした場合の合計得点と偏差値

(注) 表中の“-”は無回答であったことを示す。

(http://254index.orgより転載)

大学入試問題の分類

factoid型質問

- Factoid (事実型)
- Slot-Filling (穴埋め)
- True/False (相対的正誤判断)
- True/False Combo (絶対的正誤判断)

yes/no型質問

- Time (時系列)
- Graph (グラフ)

- Essay (論述)
 - Complex Essay (大論述)
 - 指定語句あり
 - 指定語句なし
 - Simple Essay (小論述)
 - 指定語句なし

non-factoid型質問

14

NTCIR

Factoid (事実型)

問 3 下線部③に関連して、キューバ危機が起こったときのアメリカ合衆国大統領の名として正しいものを、次の①～④のうちから一つ選べ。 12

① トルーマン
② アイゼンハワー
③ ケネディ
④ ジョンソン

選択肢は基本的に同じ質問タイプ

選択肢がある場合、質問タイプの重要性低

15

NTCIR

Slot-Filling (穴埋め)

問 4 下線部③の時期に大西洋で行われた交易について述べた次の文中の空欄 ア ～ ウ に入れる語の組合せとして最も適当なものを、下の①～④のうちから一つ選べ。 4

・文内の複数箇所が欠落している場合がある

18世紀の大西洋では、西ヨーロッパからの ア などを、西アフリカにおいて イ に交換し、これを西インド諸島の植民地などに運び、そこから ウ などをヨーロッパに送る三角貿易が行われていた。

① アー砂糖 イー火器 ウー奴隷
② アー砂糖 イー奴隷 ウー火器
③ アー火器 イー砂糖 ウー奴隷
④ アー火器 イー奴隷 ウー砂糖
⑤ アー奴隷 イー火器 ウー砂糖
⑥ アー奴隷 イー砂糖 ウー火器

難度:
選択肢がある場合
Slot-Filling < Factoid
選択肢がない場合
Slot-Filling > Factoid

16

NTCIR

True/False (相対的正誤判断)

問 6 下線部⑥に関連して、イスラム王朝・国家について述べた文として正しいものを、次の①～④のうちから一つ選べ。 6

① 13世紀に、アウラングゼーブが北インドに奴隷王朝を建てた。
② 15世紀に、西アフリカではソングアイ王国に替わってマリ王国が興った。
③ 16世紀に、オスマン朝がアナトリア(小アジア)に興った。
④ 20世紀に、イラン革命の結果、ホメイニがイランの指導者となった。

・選択肢の中からもっとも正しい(間違っている)ものを選択

17

NTCIR

True/False Combo (絶対的正誤判断)

問 6 下線部⑤に関連して、世界史上見られた広域的な労働力の流れについて述べた次の文 a と b の正誤の組合せとして正しいものを、下の①～④のうちから一つ選べ。 24

a 東南アジアがヨーロッパ諸国に植民地支配されるに及んで、中国から移住し労働に従事する者がとれた。
b ヨーロッパ諸国の中には、19世紀に、奴隷貿易を禁止する動きがあった。

・各文単独で正誤判断する必要がある

① aー正 bー正 ② aー正 bー誤
③ aー誤 bー正 ④ aー誤 bー誤

難度: True/False Combo > True/False

18

NTCIR QA Lab

Time(時系列)

問 5 下線部④に関連して、次の年表に示したa～dの時期のうち、インドの民族運動の弾圧をねらったローラット法が公布された時期として正しいものを、下の①～④のうちから一つ選べ。 **5**

・年表などの知識が必要

a
1885 年 インド国民会議の開催
b
1906 年 国民会議派、カルカッタ大会を開催
c
1930 年 第2次サティヤーグラハ(非暴力・不服従の抵抗運動)
d

① a ② b ③ c ④ d

19

NTCIR QA Lab

Graph(グラフ)

問 8 次の地図中の地域a～dのうち、下線部⑤で「私たちの国」と呼ばれている地域を示すものとして最も適当なものを、下の①～④のうちから一つ選べ。

17

・画像認識などの処理が必要

① a ② b ③ c ④ d

20

NTCIR QA Lab

Complex-Essay(大論述)

第 1 問

・複数の観点から述べる

次の文章は日本国憲法第二十条である。

第二十条 信教の自由は、何人に対してもこれを保障する。いかなる宗教団体も、国から特権を受け、又は政治上の権力を行使してはならない。

2. 何人も、宗教上の行為、祝典、儀式又は行事に参加することを強制されない。

3. 国及びその機関は、宗教教育その他いかなる宗教的活動もしてはならない。

この条文に見られるような政治と宗教の関係についての考えは、18 世紀後半以降、アメリカやフランスにおける革命を経て、しだいに世界の多くの国々で力をもつようになった。

それ以前の時期、世界各地の政治権力は、その支配領域内の宗教・宗派とそれらに属する人々をどのように取り扱っていたか。18 世紀前半までの西ヨーロッパ、西アジア、東アジアにおける具体的な事例を挙げ、この3つの地域の特徴を比較して、解答欄(イ)に30 行以内で論じなさい。その際に、次の7つの語句を必ず一度は用い、その語句に下線を付しなさい。

具体的事例

特徴を比較

ジズヤ 首長法 グラフマ ナントの王令廃止
ミット 現藩院 親邦教会制

21

NTCIR QA Lab

Simple-Essay(小論述)

問② 中国においては、新石器時代以来、城壁都市が建設され、やがて君主をいただく国となった。そうした国々を従えた大国のいくつかは、王朝として知られている。以下の(a)・(b)の間に、冒頭に(a)・(b)を付して答えなさい。

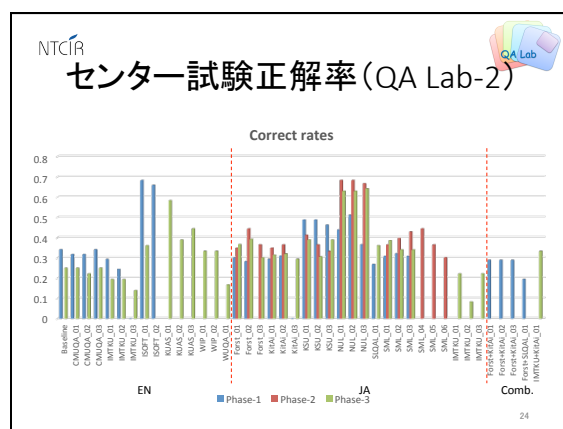
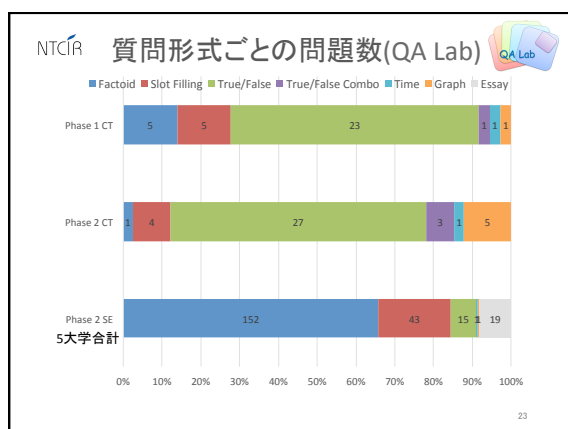
(a) 最古とされる王朝の遺跡が20 世紀初頭に発掘された。そこで出土した記録は、王朝の政治がどう行われたかを証言している。その政治の特徴を2 行以内で説明しなさい。

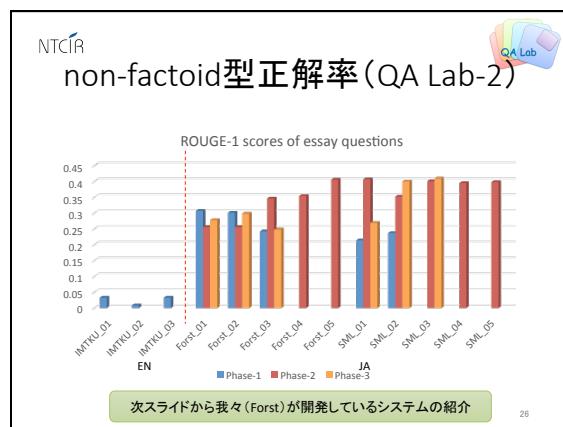
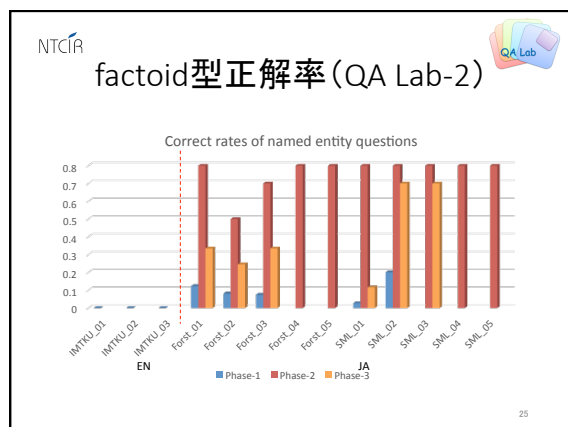
政治の特徴

(b) その後、紀元前 11 世紀に華北に勢力をのびた別の王朝は、首都の移転により時代区分がなされる。移転前と移転後の首都名を挙げ、移転にともなう政治的変化を2 行以内で説明しなさい。

・1つの観点から述べる

22





語句記述問題

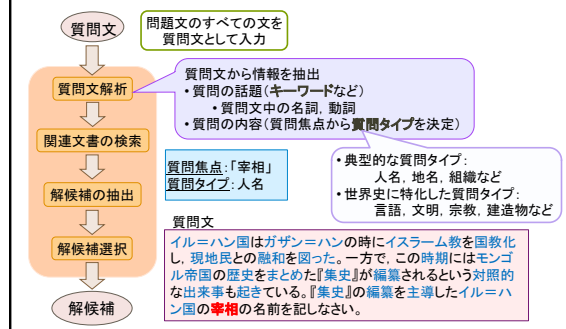
factoid型質問

問題例

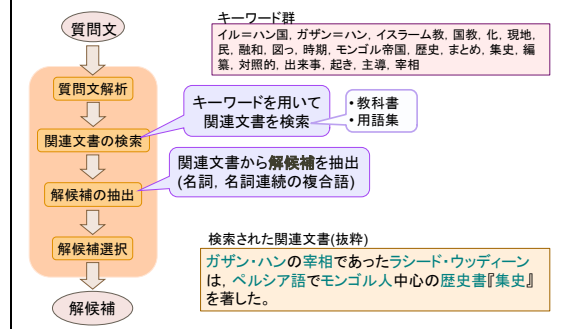
- イル＝ハン国はガゼン＝ハンの時にイスラーム教を国教化し、現地民との融和を図った。一方で、この時期にはモンゴル帝国の歴史をまとめた『集史』が編纂されるという対照的な出来事も起きている。『集史』の編纂を主導したイル＝ハン国の宰相の名前を記しなさい。
- 解答:「ラシード・ウッディーン」

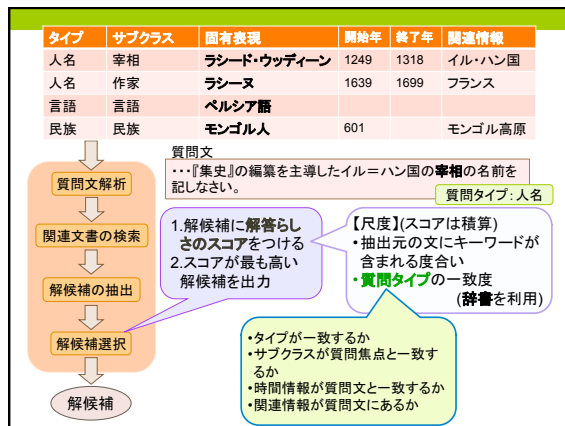
- 問題文は複数文から成る
- 文章ではなく語句での解答

語句記述問題解答システムの概要



語句記述問題解答システムの概要





結果と今後の課題

- QA Lab-2 Phase-2では8/10点
 - ただし、Phaseごとのバラつきが大きく安定した結果を残せていない

- 今後の課題
 - 知識源の不備
 - タイプ-サブクラスの不備
 - 「皇帝の父の名」のような質問



論述問題の種類

小論述問題 (2005年東京大学前期試験世界史大問2問1)

問題文

問1) オリエンタルあるいは西アジアに流通したヘレニズム文明は、さらにインドにも影響を及ぼしている。とりわけ、1世紀頃から南インドにおいてヘレニズムの影響を受けながら発達した美術には注目すべきものがある。その美術的特質について、3行以内で説明しなさい。

大論述問題 (2007年東京大学前期試験世界史大問1)

問題文

古来、世界の大多数の地域で、農業は人間の生命維持のために基礎食糧を提供してきた。それゆえ、農業生産の変動は、人口の増減と密接に連動した。耕地の拡大、農法の改良、新作物の伝播などは、人口成長の前提をなすと同時に、やがて商品作物栽培や工業化を促し、分業発展と経済成長の原動力にもなった。しかしその反面、凶作による飢饉は、世界各地にたびたび危機をもたらした。

以上の論点をふまえて、ほぼ11世紀から19世紀までに生じた農業生産の変化とその意義を述べなさい。解答は解答欄(イ)に17行以内で記入し、下記の8つの語句を必ず一回は用いたうえで、その語句の部分に下線を付しなさい。

指定語句

湖広熟すれば天下足る アイルランド トウモロコシ 農業革命
穀物法廃止 三圃制 アンデス 占城稻

大論述問題と模範解答

2007年東京大学前期試験世界史大問1

古来、世界の大多数の地域で、農業は人間の生命維持のために基礎食糧を提供してきた。それゆえ、農業生産の変動は、人口の増減と密接に連動した。耕地の拡大、農法の改良、新作物の伝播などは、人口成長の前提をなすと同時に、やがて商品作物栽培や工業化を促し、分業発展と経済成長の原動力にもなった。しかしその反面、凶作による飢饉は、世界各地にたびたび危機をもたらした。

以上の論点をふまえて、ほぼ11世紀から19世紀までに生じた農業生産の変化とその意義を述べなさい。解答は解答欄(イ)に17行以内で記入し、下記の8つの語句を必ず一回は用いたうえで、その語句の部分に下線を付しなさい。

Keywords

湖広熟すれば天下足る アイルランド トウモロコシ 農業革命
穀物法廃止 三圃制 アンデス 占城稻

中世ヨーロッパでは、11世紀頃から普及した三圃制農法が穀・水車の改良とあいまって農業生産力を増大させ、人口増加とそれに伴う都市の発展や東方植民の原因となった。同時期の中国では、宋代にベトナムから日照りに強い占城稻が導入され、長江下流の水稲地帯で集約的農法が発展したが、明代になると長江下流では家内制手工業が盛んとなり、原料である綿花や桑の栽培が普及したため、米の主産地は「湖広熟すれば天下足る」といわれるように長江中流域に移動した。16世紀以降にはアメリカ大陸のトウモロコシが家畜の飼料として世界各地に普及し、アンデス原産のジャガイモは主にヨーロッパの寒冷地に拡大してその人口増を支えたがアイルランドでは19世紀半ばの「ジャガイモ飢饉」によって多くの餓死者とアメリカへの移民を生むことになった。一方、18世紀のイギリスではノーフォーク農法や第2次圈い込みに代表される農業革命によって農業の資本主義化が進み、この結果土地を失った農民は都市に流入して産業革命を支える労働力となった。また産業革命によって台頭した産業資本家の要求により、地主保護のために制定されていた穀物法廃止が実現し、イギリスでは自由貿易体制が確立した。

指定語句あり論述問題と模範解答

質問の内容 「農業生産の変化とその意義を述べなさい」

直接関係ないもの

指定語句:

湖広熟すれば天下足る アイルランド トウモロコシ 農業革命
穀物法廃止 三圃制 アンデス 占城稻

アンデス原産のジャガイモは主にヨーロッパの寒冷地に拡大してその人口増を支えた

質問内容にあった使われ方をしている記述を見つける

指定語句あり論述 問題と模範解答 2007年東京大学前期試験世界史大問1

古来、世界の大多数の地域で、農業は人間の生命維持のために基礎食糧を提供してきた。それゆえ、農業生産の改良、新作物の伝播などは、人口成長の前提をなすと同時に、やがて商品作物栽培や工業化を促し、分業発展と経済成長の原動力に繋がった。世界各地にたまたま危機をもたらしたものの、農業生産の変化とその意義をふまえて、ほぼ11世紀から19世紀までに生じた農業革命の変化と、その意義を述べなさい。解答は解答欄(イ)に17行以内で記入し、下記の8つの語句を必ず一回は用いたうえで、その語句の部分に下線を付しなさい。

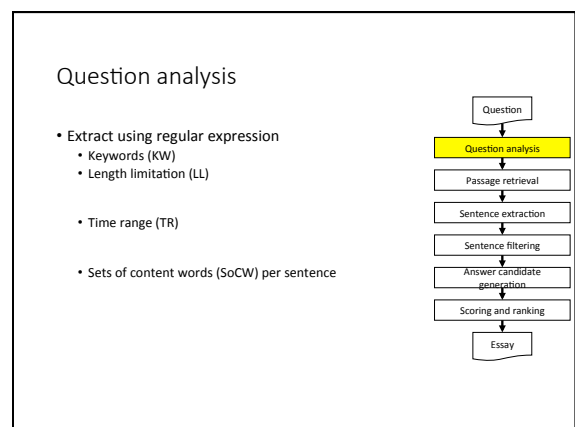
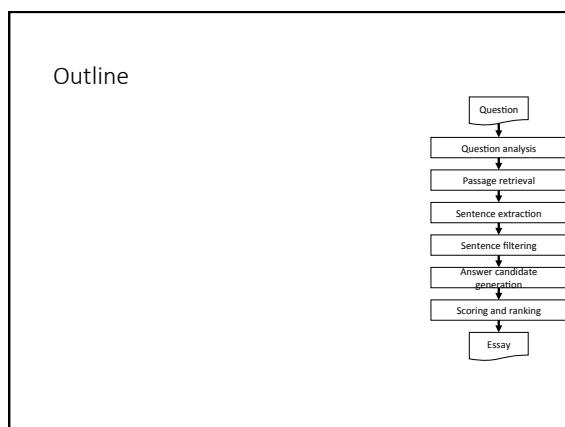
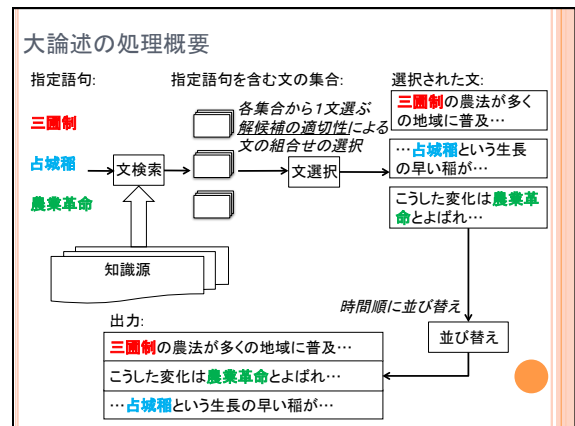
指定語句数: 8つ
制限の510文字÷8文(指定語句を含む文)
=1文あたり平均63.75字 ⇒ 短い

記述の長さ

字数制限を越えないようにしながら回答を生成

シが家畜の飼料として世界各地に普及し、アフリカ原産のジャガイモは主にヨーロッパの寒冷地に拡大してその人口増を助けた。また、アメリカへの移民を促した。代表的な農業革命は、11世紀から19世紀までに生じた農業革命の変化と、その意義を述べなさい。解答は解答欄(イ)に17行以内で記入し、下記の8つの語句を必ず一回は用いたうえで、その語句の部分に下線を付しなさい。

湖広熟すれば天下足る
 穀物法廃止
 アイルランド
 トウモロコシ
 トウモロコシ
 アンデス
 農業革命
 占城稲



大論述の制約

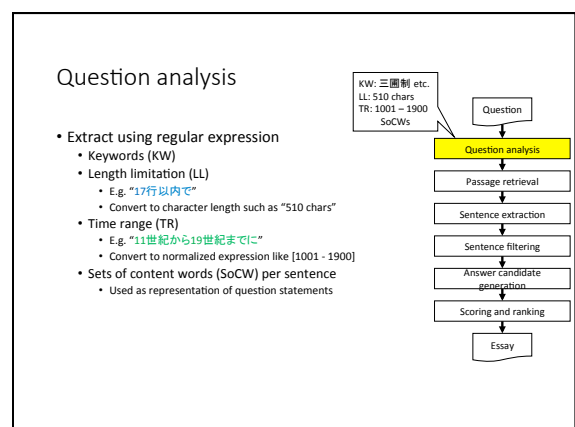
2007年東京大学前期試験世界史大問1

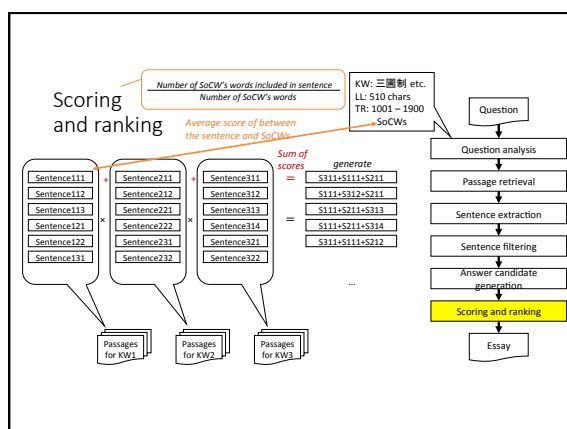
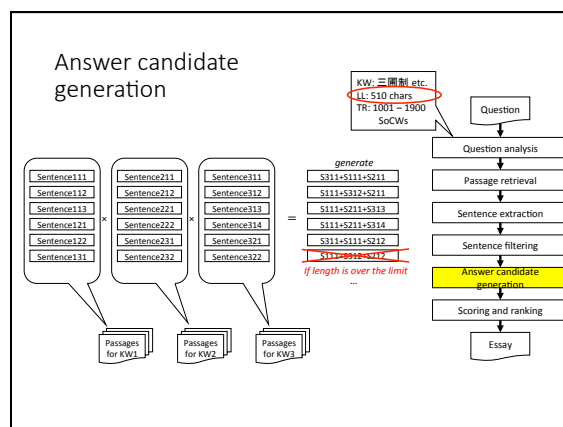
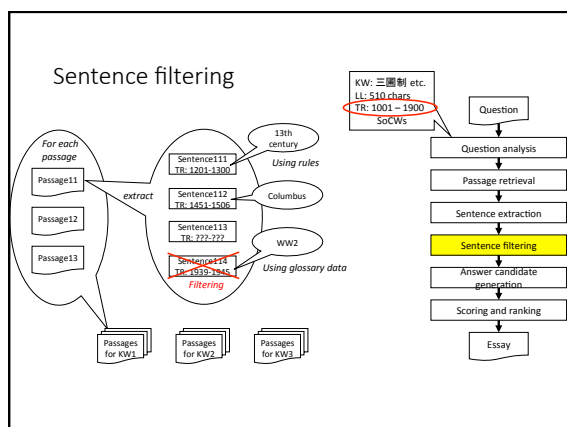
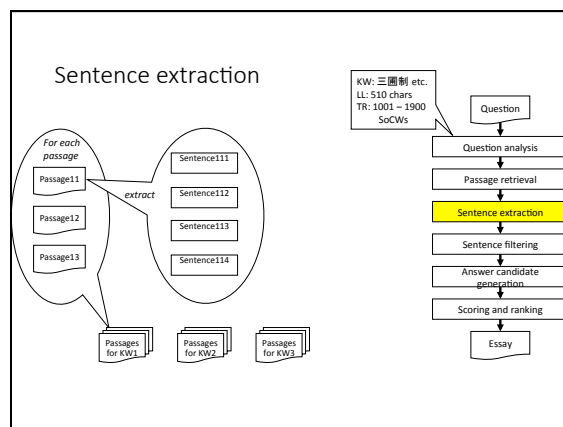
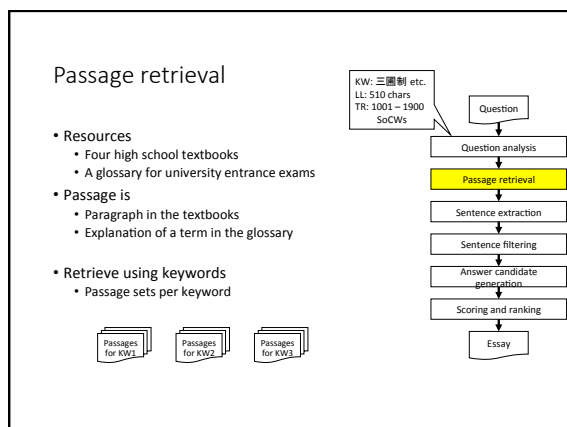
古来、世界の大多数の地域で、農業は人間の生命維持のために基礎食糧を提供してきた。それゆえ、農業生産の変動は、人口の増減と密接に連動した。耕地の拡大、農法の改良、新作物の伝播などは、人口成長の前提をなすと同時に、やがて商品作物栽培や工業化を促し、分業発展と経済成長の原動力に繋がった。世界各地にたまたま危機をもたらしたものの、農業生産の変化とその意義をふまえて、ほぼ11世紀から19世紀までに生じた農業革命の変化と、その意義を述べなさい。解答は解答欄(イ)に17行以内で記入し、下記の8つの語句を必ず一回は用いたうえで、その語句の部分に下線を付しなさい。

Length limit
 (1 line = 30 characters)

Time range
 11世紀から19世紀までに

Keywords
 湖広熟すれば天下足る
 穀物法廃止
 アイルランド
 トウモロコシ
 トウモロコシ
 アンデス
 農業革命
 占城稲





大論述の出力結果 (2007年東京大学世界史大問1)

この時代はおおむね気候が温和で、**三國制**の普及や犁・水車の改良など農業技術の進歩により農業生産は増大し、人口も飛躍的にふえた。**トウモロコシ**のほかにジャガイモ、トマト、カボチャ、カカオ豆などが栽培され、のちに世界各地に伝えられることになった。明末には長江中流域の湖広(現在の湖北・湖南省)が新たな穀倉地帯となり、「**湖広熟すれば天下足る**」と称せられた。11世紀ころ、**古城種**という生長の早い稲がヴェトナム方面から伝えられ、台地や海岸部の水利の悪い土地に栽培されるようになった。ペルーの**アンデス**山地でも高度な文明が生まれていたが、14世紀なかば、ペルー南部高原のインカ族がクスコを都として勢力を強め、15世紀には多くの国を統一して大帝國となった。経済と社会の産業化ヨーロッパの人口増大と**農業革命**イギリス・オランダ・北フランスなど北西ヨーロッパを先頭にして、西ヨーロッパでは、18世紀前半には休耕地を設けずに輪作など新農法が普及して農業生産力が増大し、食糧事情は好転した。そのイングランドが周辺のケルト系のウェールズやスコットランド、**アイルランド**などを併合や征服することによって、1801年以来、現在の連合王国になったのである。

Result

• NTCIR-12 QA Lab-2 task

• Used data

- University of Tokyo from 2000 to 2011
- Sundai mock exam on Aug, 2016

• Evaluation

- ROUGE-1 score (using model answer as reference)
- Human expert's marks (only Sundai mock exam)

	Phase-1	Phase-2	Phase-3
ROUGE-1 score	0.546	0.541	0.592
Human expert's marks	-	9/26	-

The average of high school students' marks was 4.3

今後の課題

- 言い換えや要約などの活用
 - 600字÷25項目＝24字
 - Ex. 「395年、ローマ帝国は東ローマ帝国と西ローマ帝国に分裂した。その結果...」⇒「東西分裂後...」
- 「正しいが重要ではない」記述
 - 質問の意図(文脈)の理解 ⇒ 適切な回答
- 知識源の信憑性
 - 現在では誤りとされる説、用いられない語句
- 自動評価手法の確立

NTCIR-13 QA Lab-3への招待



Question and Answer Data



- Training Data
 - 8 years of National Center Test (JA and EN)
 - 5 years of Second-stage Examinations for 5 univs (JA and EN)
 - 11 years of Essay questions of 2 univs' entrance exams (JA)
 - **Nugget data** for 11 years of Essay questions of 2 univs' entrance exams (JA)
 - 9 years of Mock Exams for National Center Test (JA)
 - 5 years of Mock Exams for the University of Tokyo entrance exam (JA)
- Test Data
 - 2 sets (JA and EN)

From QA Lab-1 and -2 training and test data



Knowledge Sources



- Four Japanese high school textbooks
 - Available in Japanese
- A snapshot of Wikipedia
 - Available in English and Japanese
- World History Ontology
 - Available in Japanese
 - <http://researchmap.jp/zoeai/event-ontology-EVT/>
- Participants are free to use any resources



Tools



- Two Baseline Systems for Center Test (Japanese and English)
 - UIMA module-based end-to-end QA systems
- **One Baseline System for Essay Question (Japanese)**
- One Passage Retrieval Module
 - to enhance the module-based collaboration.
- Scorer for National Center Test
- Format Checker for National Center Test and Second-stage Examination

