

合成可能性評価に向けた クラウドソーシングの利用 (Toward Crowdsourcing Evaluation of Synthetic Accessibility)

馬場 雪乃（東大）、
磯村 哲（地球快適化インスティテュート）、
鹿島 久嗣（東大）、中野 浩史（東大）
CBI学会2013年大会、2013年10月31日

概要：クラウドソーシングによる 化合物合成可能性判定の実現可能性を検討

- **大量の**化合物の合成可能性を人間に判定させたい
- **クラウドソーシング**を利用し
「**専門家以外の人も含む**多数の人々が、それぞれ
短時間で合成可能性を判定する」仕組みを目指す
- 実現可能性検討のため初期実験で以下を調査
 1. 専門家は短時間でも合成可能性を判定することができるのか
 2. 専門家ではない人が合成可能性を判定することができるのか

背景：合成可能性の自動判定手法は未確立、 専門家による判定は人手不足

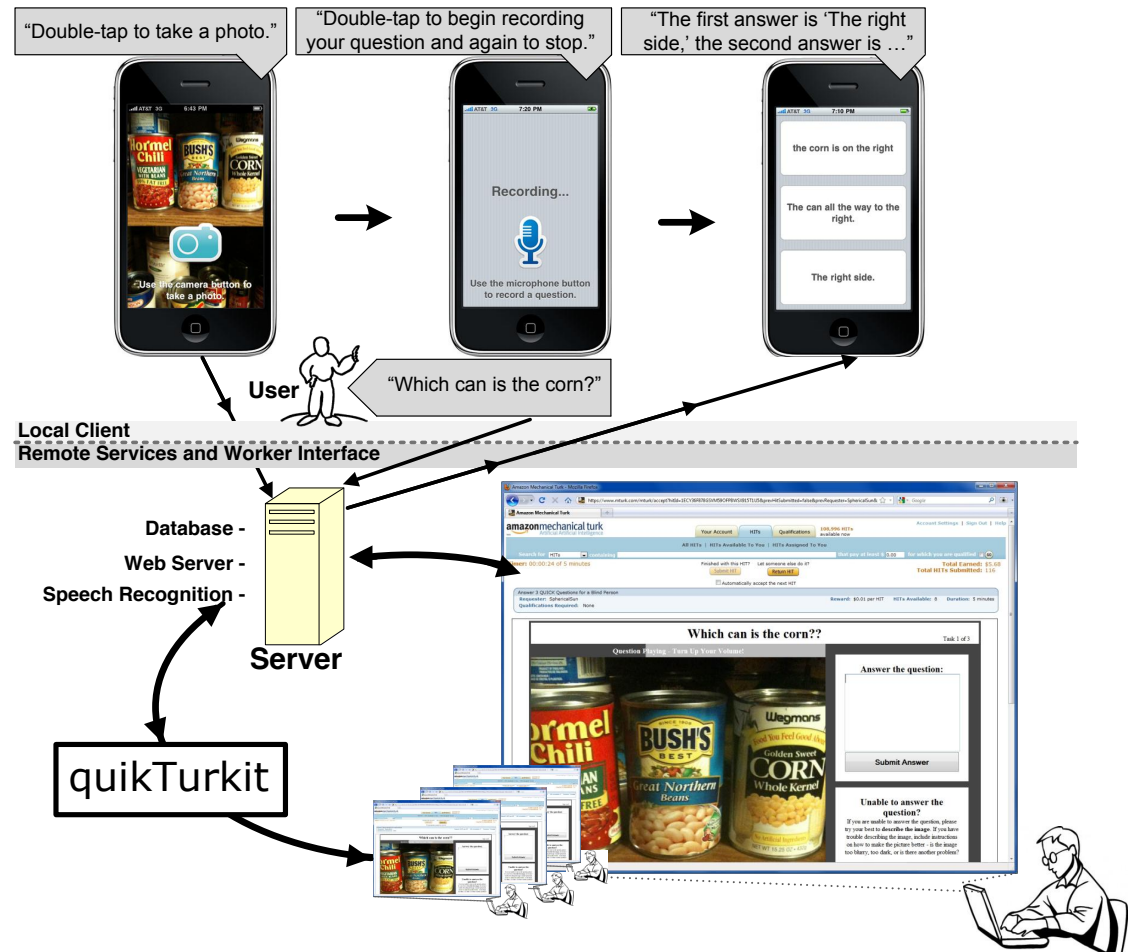
- 分子設計システムを用いた創薬では、
合成可能性による化合物のスクリーニングが必要
- 自動判定手法はまだ精度・計算時間の面で実用レベルには達していない
 - Molecular complexityにもとづく手法
 - 合成経路設計システム
- 専門家による判定が現実的な手法
 - 大量の化合物を判定させるには人手不足

クラウドソーシング：大量の細切れタスクを 不特定多数の人々に発注できる電子市場

- クラウドソーシング：
「インターネットを通じて不特定多数の人々に仕事を依頼する仕組み」
電子市場の例：Amazon社のMechanical Turk
- 「マイクロタスク」と呼ばれる、1回数秒～数分程度の規模の仕事に分割し依頼する形式が主流
- 計算機科学の分野では、データ収集・評価実験・人間を組み込んだアプリケーションの開発などで広く活用されている

人間を組み込んだアプリケーションの例： 視覚障がい者支援システムVizWiz

- 視覚障がい者が写真と音声による質問文を送る
- VizWizはクラウドソーシングを用いて、短時間で回答を返す



Bigham et al.: VizWiz: nearly real-time answers to visual questions, UIST 10, 2010.

※図は元論文Figure 1より引用

我々のビジョン：クラウドソーシングを利用して 創薬プロセスに人間を効率的に組み込む

- 人間を組み込んだ合成可能性判定システムを実現、
創薬プロセスの部品にする
 - 合成可能性を判定したい化合物が
システムに送られる
 - システムはクラウドソーシングを用いて
短時間で回答を返す
- さらには、蓄積されるデータを用いて
機械による合成可能性判定の精度向上を目指す

クラウドソーシングによる合成可能性判定の実現可能性検討のため初期実験を実施

- 初期実験で検討したい項目
 1. 専門家は短時間でも合成可能性を判定することができるのか
 2. 専門家以外の人々に合成可能性を判定させることができるのか

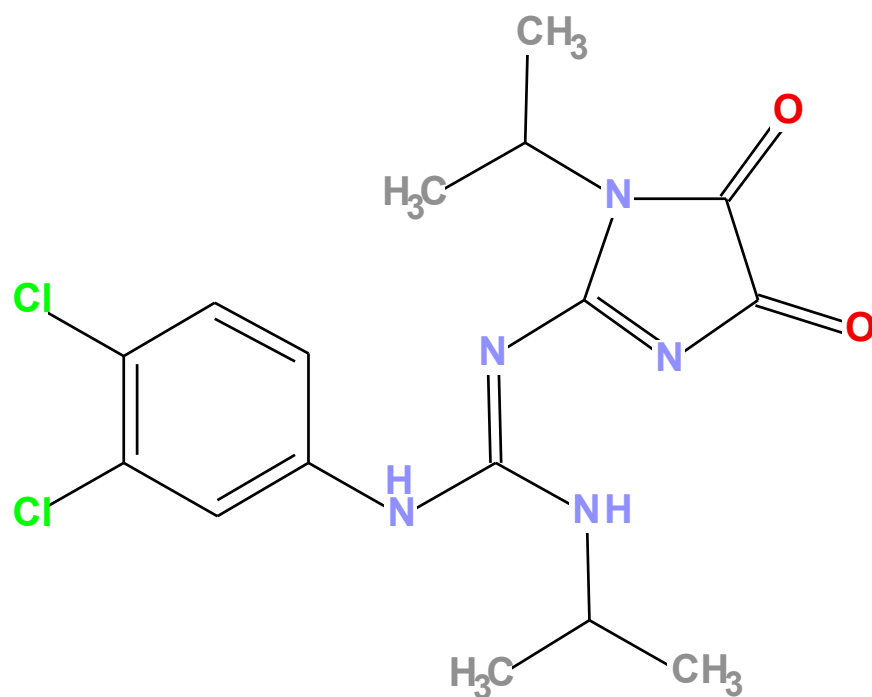
初期実験参加者：

専門家・準専門家・非専門家

- 専門性の異なる人々にいくつかの化合物の合成可能性判定を依頼し結果を比較
 - **専門家** 4名：
製薬会社に勤務する現役ケミスト
 - **準専門家** 4名：
他分野のケミスト・現役から離れたケミスト等
 - **非専門家** 100名：
商用クラウドソーシングサービスLancers上で雇用した作業者

実験インターフェース（専門家、準専門家向け）

1 問めの作業（30問中）



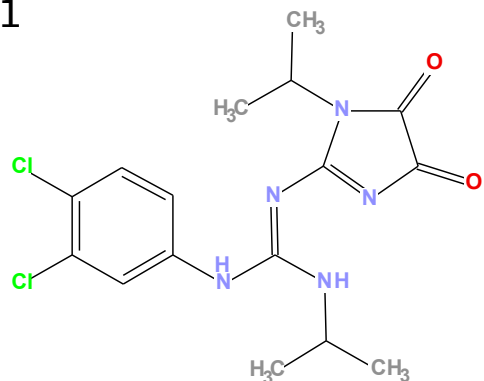
- 1（難しい）
- 2（やや難しい）
- 3（普通）
- 4（やや易しい）
- 5（易しい）

※一つの化合物についての回答を送信すると、次の化合物が出題される
※一化合物あたり「数十秒～数分程度で回答」するように指示

実験インターフェース (非専門家向け)

表示される化合物について、合成難易度 (= 作るのがどれくらい難しいか) を直感で判断してください。間違っても構いません。

化合物 1



- 1 (難しそう)
- 2 (やや難しそう)
- 3 (どちらともいえない)
- 4 (やや簡単そう)
- 5 (簡単そう)

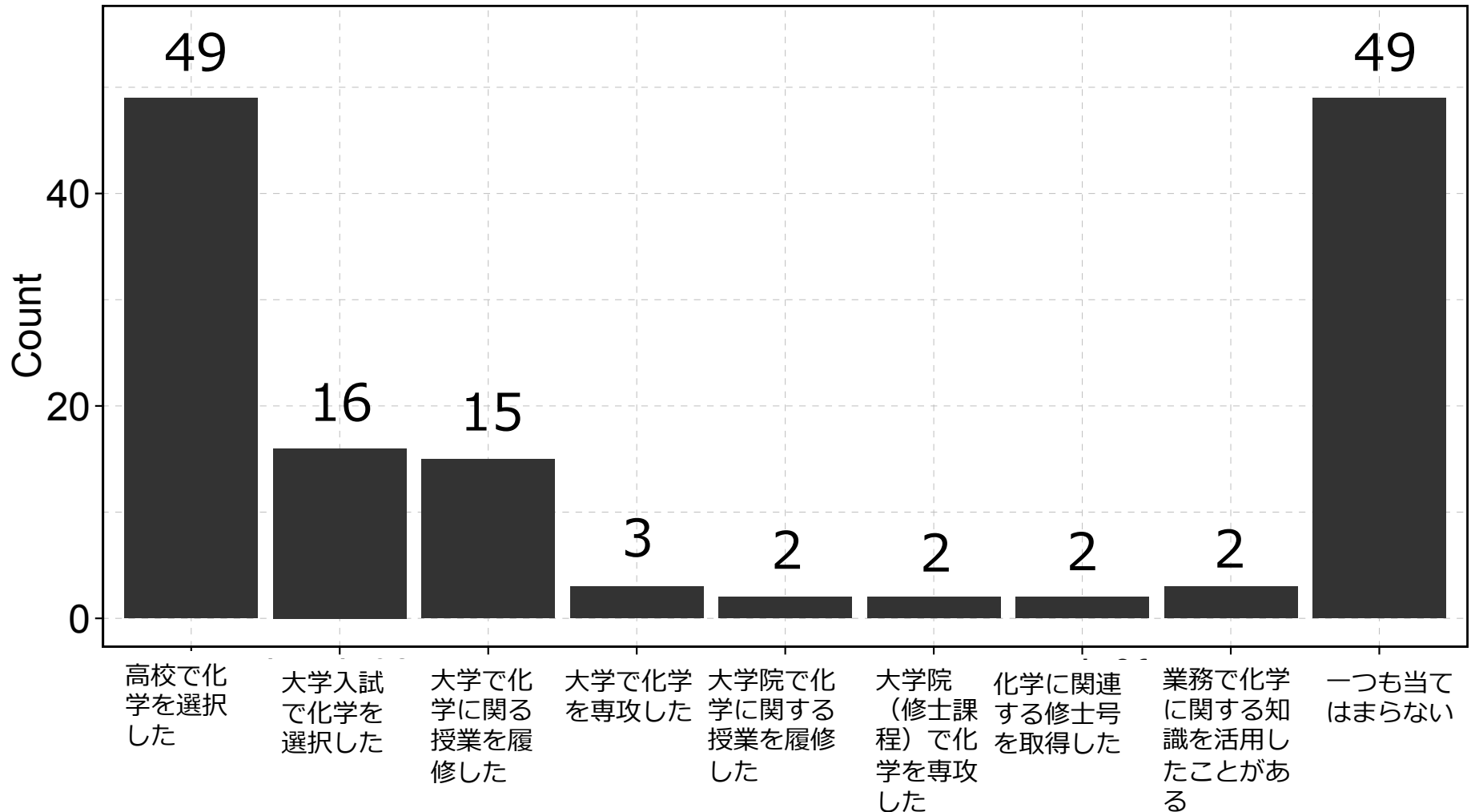
化合物 2

⋮

※一度に全ての化合物について回答するよう依頼

非専門家の知識レベル:

中高で化学を学んだ程度の人が大多数



※当てはまるものを全てを挙げるよう指示

既存化合物、合成困難化合物、新規化合物の 3種類の問題セットを用意

- **既存化合物 (a) 10個**
 - 既存のセットからランダムに選択
- **合成困難化合物 (a') 10個**
 - 既存化合物 (a)の構造を手動で一部変化させ、合成困難にしたもの
- **新規化合物 (b) 10個**
 - 専門家にとっても未知の化合物
 - 既存化合物(a)の類似構造をGDB-17から1000個程度取得
 - そのうち、化合物DB内の環構造をもたない化合物を抽出

本実験において、専門家は極めて短い時間で化合物の合成可能性を判定

	専門家 A	専門家 B	専門家 C	専門家 D	準専門家 家E	準専門家 家F	準専門家 家G	準専門家 家H
各回答者の 平均時間 (秒)	17.3	14.5	29.7	28.3	22.3	52.7	151.6	15.4
	専門家平均：22.5				準専門家平均：60.5			

- 専門家は一化合物あたり平均22.5秒で回答
- 準専門家は専門家よりも回答に時間を掛けていた

※回答時間は、「合成可能性」の回答だけではなく「有用性」への回答とコメント（任意）の記入時間を含む

評価方法：「合成困難化合物」の方が「既存化合物」より難易度が高いと、正しく判断できるか？

● 評価 (1)

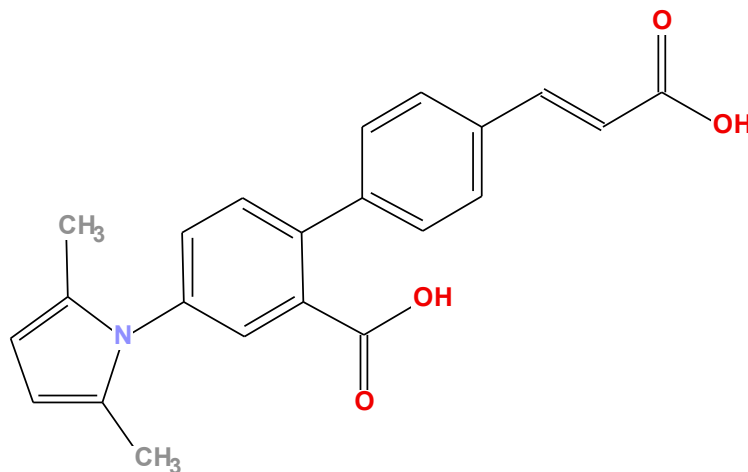
- 「既存化合物 (a)」と「合成困難化合物 (a')」では、合成困難化合物の方が合成が難しいはず。正しく判断できているか？

● 評価 (2)

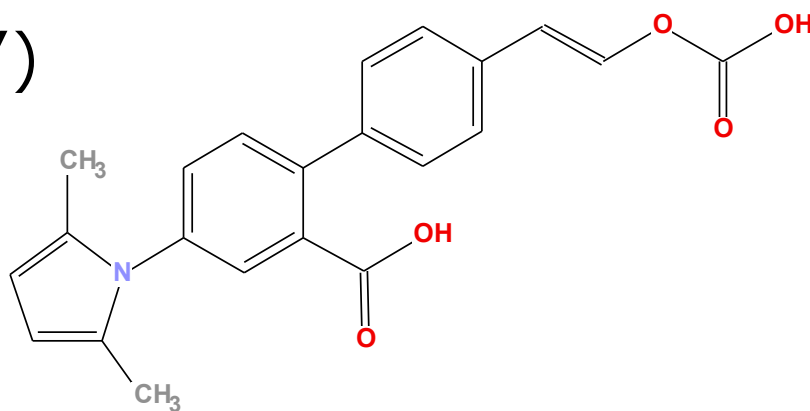
- 準専門家・非専門家は専門家と似た判断ができているか？

評価 (1) の例

既存化合物 (a)



合成困難化合物 (a')



十分な判定能力があれば「(a)より(a')の方が合成が難しい」と答えられるはず

評価（1）の結果：

専門家は短い回答時間でも正しく判断できた

専門家の正解率	
専門家A	1.00
専門家B	0.95
専門家C	0.80
専門家D	0.90
平均正解率	0.91

準専門家の正解率	
準専門家E	0.45
準専門家F	0.65
準専門家G	0.55
準専門家H	0.75
平均正解率	0.60

非専門家の正解率	
平均正解率	0.52

- 専門家は0.91と高い平均正解率
- 平均正解率は、専門家 > 準専門家 > 非専門家

※ 「(a)の難しさ<(a')の難しさ」と正しく回答した場合1点、タイの場合0.5点与えるものとし、10ペアでの平均得点を正解率とした

評価方法：専門家と似た判断ができていますか？

- 評価 (1)

- 「実在化合物 (a)」と「合成困難化合物 (a')」
では、合成困難化合物の方が合成可能性が低い
はず。正しく判断できているか？

- 評価 (2)

- 準専門家・非専門家は専門家と似た判断ができ
ているか？

評価 (2)の結果：準専門家は、 非専門家よりも専門家と似た回答ができた

	既存化合物 (a)	合成困難 化合物 (a')	新規化合物 (b)	平均
専門家間の 距離の平均	0.954	0.619	0.789	0.787
「専門家、準専門家」 間の距離の平均	1.440	1.071	0.951	1.154
「専門家、非専門家」 間の距離平均	1.515	1.092	1.331	1.313

- 専門家との回答の類似性は、
準専門家 > 非専門家

新規化合物においても
合成困難化合物と同程
度の高い合意を確認

※各人の回答を「平均0, 分散1」になるよう正規化した上で
「回答間のユークリッド距離/ $\sqrt{\text{問題数}}$ 」を「距離」として計算

まとめと今後の予定

- まとめ
 - クラウドソーシングによる合成可能性判定の実用性を検討するため初期実験を実施
 - 専門家は一化合物あたり20秒程度でも正しく合成可能性を判定できた
 - 非専門家にとっては合成可能性の判定は難しい。準専門家は、非専門家よりは正しく判定できた
- 今後の課題
 - 実験規模拡大
 - 準専門家を有効活用する方法の検討