

JSAI-2020チュートリアル講演 9:00-10:40
<https://www.ai-gakkai.or.jp/jsai2020/tutorial>



AI系トップカンファレンスへの 論文採択に向けた試験対策

東北大学 大学院情報科学研究科
鈴木 潤
2020/06/10

本講演の目的

- 論文執筆
 - ツライ，時間がかかる
- 論文執筆負荷を下げる
 - 知見の共有
 - ポイントをおさえて効率的に

想定聴衆

● 対象

- これから投稿していきたいと思っている人
- 既に国際会議に投稿(and 採録)しているが査読プロセスの全体像があまり把握できていない人
- さらに効率よく論文を書きたいと思っている人

● 対象外

- 何度も「採録あり」な人
- 何度も reviewer になっている人
- 何度も「採録あり」な人が身近（共著者）にいる人
- トップ国際会議に論文を通すことに興味のない人

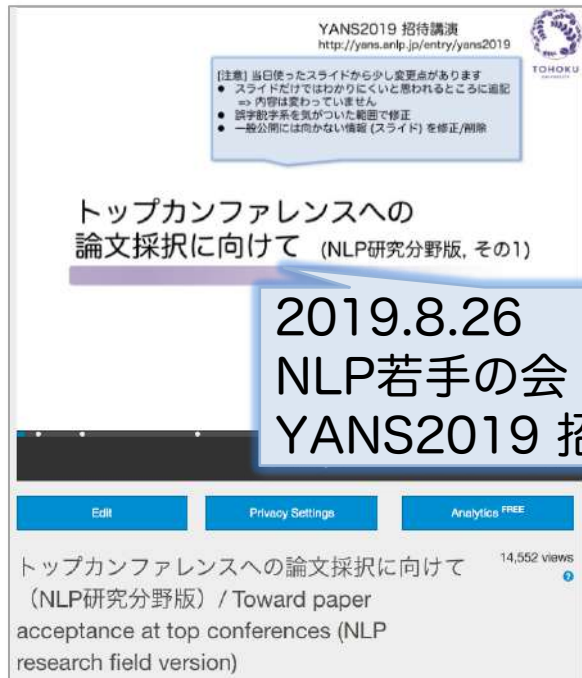
AI系トップ国際会議

- 今回の講演の(気持ちベースの) ターゲット
 - AAAI Since 1980
 - IJCAI Since 1969

 - NeurIPS Since 1986
 - ICML Since 1980
 - ICLR Since 2013
- いわゆる最難関国際会議

[参考]

- (似た内容で) 4回目の講演. . .
 - 今回は「試験対策」に特化
 - 現状分析 / 研究の見つけ方 などは以下を参考



YANS2019 招待講演
http://yans.anlp.jp/entry/yans2019

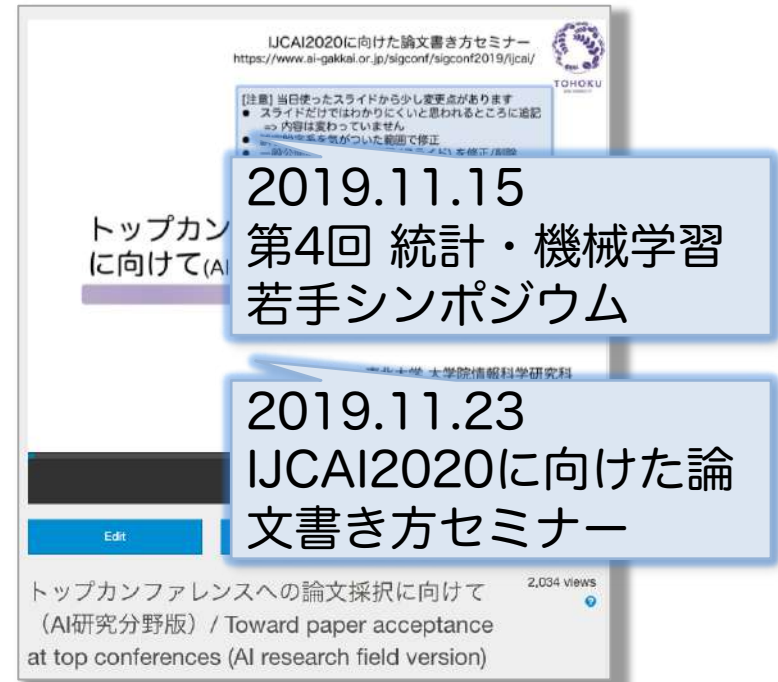
[注意] 当日使ったスライドから少し変更点があります
 ● スライドだけではわかりにくいと思われるところに追記
 => 内容は変わっていません
 ● 誤字脱字系を気がついた範囲で修正
 ● 一発公開には向かない情報 (スライド) を修正/削除

トップカンファレンスへの
論文採択に向けて (NLP研究分野版, その1)

2019.8.26
NLP若手の会
YANS2019 招待講演

Edit Privacy Settings Analytics FREE

トップカンファレンスへの論文採択に向けて 14,552 views
 (NLP研究分野版) / Toward paper acceptance at top conferences (NLP research field version)



IJCAI2020に向けた論文書き方セミナー
https://www.ai-gakkai.or.jp/sigconf/sigconf2019/ijcai/

[注意] 当日使ったスライドから少し変更点があります
 ● スライドだけではわかりにくいと思われるところに追記
 => 内容は変わっていません
 ● 誤字脱字系を気がついた範囲で修正
 ● 一発公開には向かない情報 (スライド) を修正/削除

2019.11.15
第4回 統計・機械学習
若手シンポジウム

2019.11.23
IJCAI2020に向けた論
文書き方セミナー

Edit

トップカンファレンスへの論文採択に向けて 2,034 views
 (AI研究分野版) / Toward paper acceptance at top conferences (AI research field version)

<https://www.slideshare.net/JunSuzuki21/2019-0826-yansinvitedtalk>

<https://www.slideshare.net/JunSuzuki21/20191123jsaiinvitedtalk-205359389>

概要

- 可能な限り効率的に
=> 敢えて「試験対策」と考える
- 査読プロセス概要
 - ・ 誰がどうやって採否をきめてるの
- 論文執筆時の基本方針
 - ・ 何に気をつけるか
- 論文執筆時の細かいテクニック
 - ・ 見栄えをよくするには
- 賢い author response (rebuttal)の 取り組み方
 - ・ 査読者との対話

[参考] 一応のお断り

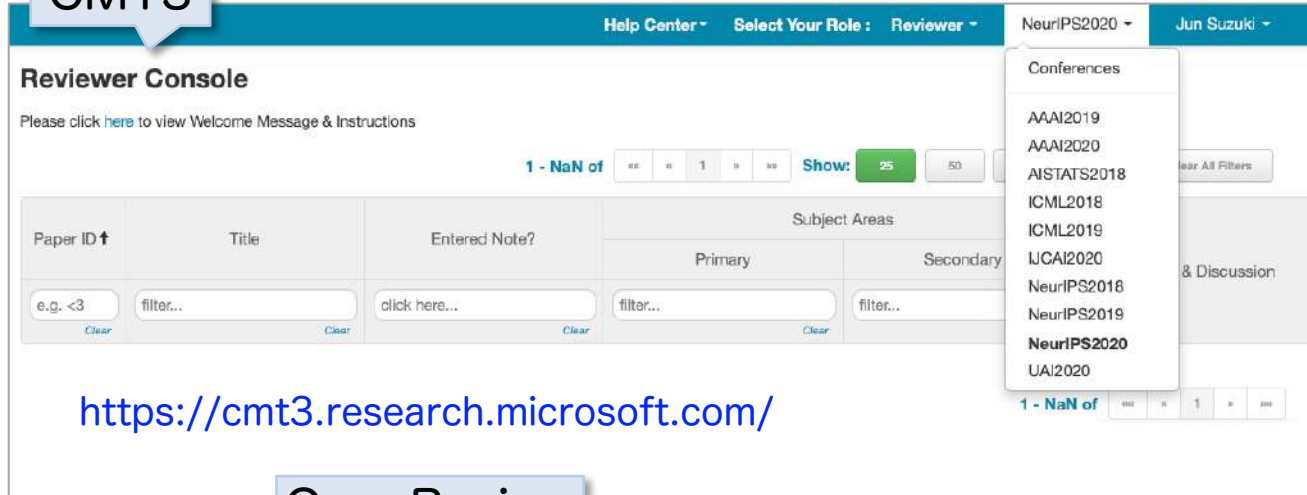
- 国際会議に採録される論文 ≠ 良い研究
 - 採録されていなくても良い研究はたくさんある
 - 国際会議に採録されることが研究の目的ではない

- 最難関国際会議に採録 => 一定の知名度
 - (特に若手にとって) 研究能力を示す良い場
 - その後のキャリアにつながる(かも)

査読プロセス

Review system

CMT3



Help Center Select Your Role : Reviewer NeurIPS2020 Jun Suzuki

Reviewer Console

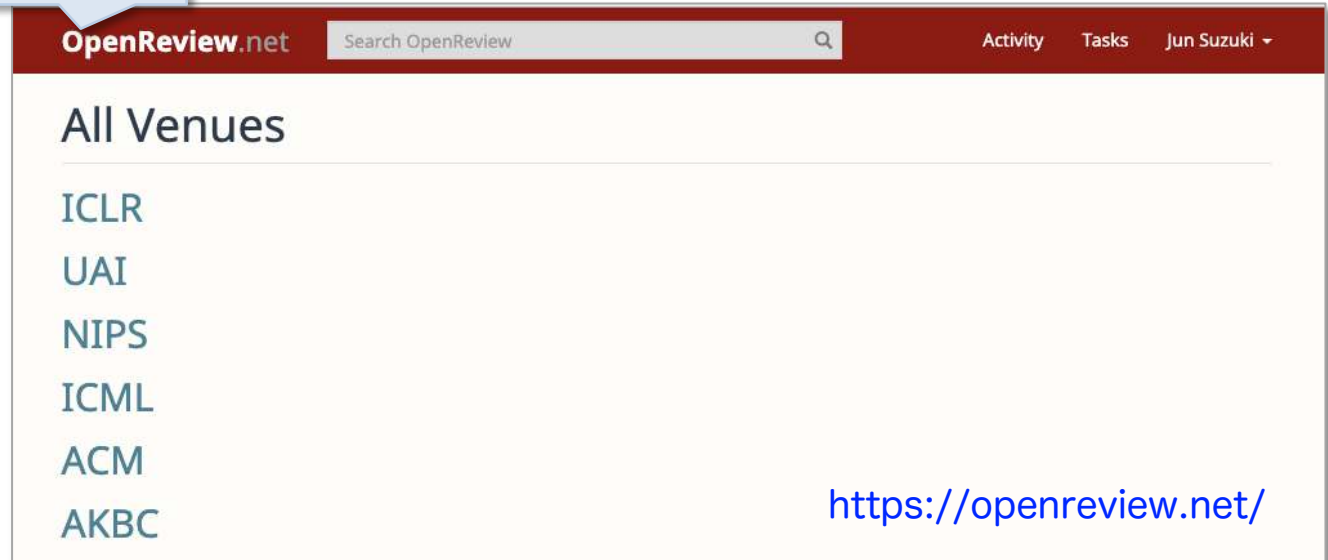
Please click [here](#) to view Welcome Message & Instructions

1 - NaN of Show: 25 50

Paper ID ↑	Title	Entered Note?	Subject Areas	
			Primary	Secondary
e.g. <3	filter..	click here...	filter..	filter..

<https://cmt3.research.microsoft.com/>

OpenReview



OpenReview.net Search OpenReview Activity Tasks Jun Suzuki

All Venues

- ICLR
- UAI
- NIPS
- ICML
- ACM
- AKBC

<https://openreview.net/>

論文の採否を決める人達

- 3 または 4階層
 - Program Chair
 - Area Chair (AC)
 - Senior Program Committee (SPC: meta-reviewer)
 - Program Committee (PC: reviewer)

[参考] ACL-2020

- 1, Program Chair
- 2, Senior Area Chair (SAC)
- 3, Area Chair (AC)
- 4, Program Committee (PC)

[参考] IJCAI-2020

- 1, Program Chair
- 2, Area Chair (AC)
- 3, Senior Program Committee (SPC)
- 4, Program Committee (PC)

名前が微妙に違う

典型的な論文査読手順

投稿論文数増加の影響か最近 review方式も様変わりしてきているので注意が必要

[著者] 論文投稿

[PC] 分野 (Area Chair) 毎に論文割り振り

[査読者] Bidding

[PC] 査読者に論文割り振り

[査読者] 査読

[著者] Author response / Rebuttal

[査読者, AC] 議論

[AC] 論文の一次評価 / ランキング

[PC] 最終決定

・登場人物

[著者] 論文を{書いた/投稿した}人

[査読者] 投稿した論文を読んで審査する人
一本の論文に最低3人程度が割り当てられる
一人の査読者に3~10本程度が割り当てられる

[AC] 同一分野の論文に対して査読結果をみて一次評価をする人/査読者のまとめ役
一人のACに数十本の論文が割り当てられる

[PC] 論文の最終的な採否を決定する人
2, 3人

査読割り振りの決定方法

● Toronto Paper Matching System (TPMS)

<https://torontopapermatching.org/webapp/profileBrowser/login/>

- + Bidding
 - YES Eager
 - Maybe Willing
 - In a pinch
 - No No
 - COI COI

今年からGoogle Scholar, Semantic Scholar, DBLP とのリンクを利用に変更?



● YES 19



● YES 0

査読時に査読者がすること

- コメントを書く
 - [Summary]
 - [STRENGTH / REASON to ACCEPT]
 - [WEAKNESS / REASON to REJECT]
 - [QUESTIONS FOR THE AUTHORS]

- スコアをつける
 - 例：10段階の場合

10, top-10%	5, Weak Reject
9, top-30%	4, Reject
8, Strong Accept	3, Strong Reject
7, Accept	2, Clear Reject
6, Weak Accept	1,

査読の現状分析

● 査読コスト

- First review: 真面目にやると1論文に最低でも1日
 - 参考文献の精査
 - Weaknessや質問の根拠を用意する
- Rebuttal対応: 期間終了まで断続的にコスト発生
 - 稀にキレられたりしてショックをうけたりする
 - Discussionのための根拠さがし

● 査読者の気持ち

- [一般論] 時間かかる, 報酬ない
 - ⇒ 積極的に頑張る気持ちは薄い (と思っておく)

論文執筆時の基本方針

論文執筆時の基本方針

- 論文は伝わらなければ意味がない
=> 読み手（査読者）の気持ちを常に考えて書く

着眼点

- 査読者 => 同分野の研究者
 - つまり，論文の採否を決めるのが同分野の研究者ということ
 - 査読者は積極的に頑張ってくれるわけではない
- 査読者 (同分野の研究者) が{喜ぶ/驚く}ことを考える
 - 手法：こんなことできるんだ，へー
 - 結果：これがあれば，自分(査読者)の研究が進む/助かる
 - 関連研究の網羅性 (<= ためになる)
 - 面白い分析結果

- 論文内容の考え方

- 採択率 20-30%

- 投稿数：5000 => 採録数 1000 !?

=> 全ての論文に勝つ，といった条件は必要ない

- 防御力が大事

- 隙のない論文，ツッコミどころが少ない論文

読み手に対する負荷

- 読み手 (査読者) の負荷が低い論文がよい
- 読み手の負荷があがる要因
 - 読み手に「察して」もらうを前提とした曖昧な記述
 - 書き手の暗黙の前提に根差した記述
 - エラーを含む数式やアルゴリズム
 - 論文の目的が曖昧
 - Introductionの内容が不明瞭
 - 「Introductionの出来が悪ければどんなに他の内容が良くても落ちる」と思っておく
 - 同じ内容でも採否が大きく変わる

フリオチ (内容の一貫性) 重要

- 何を目的とした論文なのか
 - Introductionで述べる
 - 主張(claim)/貢献(contribution)を明確に述べる
- 論文で示したい内容を示しているか？
- 実験/分析は主張/貢献をサポートしているか？
 - 仮説の回収など

主なチェック項目

- 基本的な方針：読者の負荷が高い論文はネガティブ評価！
=> 初見の人がスラスラ読めるか？が基準
- 説明なしに新概念を当たり前のよう書いていないか
 - ・ 読者は著者のもつ知識/前提条件を共有していない
- 専門用語の統一
 - ・ 同じ事象を様々な言い方で表記すると読み手は混乱
- 文単位での論理の飛躍はないか
 - ・ 文と文の繋がりを重視
- 一文で複数の事象を表現しようとしていないか
 - ・ 基本読みにくい文になる => 複数の文に分割する

論文執筆時の細かいテクニック

推敲重要

- 初稿 => 投稿版 ではない
 - これは、かなり書ける人でも落ちる可能性が高い
- なんども推敲を重ねる
 - 初稿が書き終わってから、最低でも5回、できれば10回ぐらい？
- 最後は音読する
 - 自分がスラスラ読めない論文は、他の人では絶対に読めない（高負荷），と思う

数式(notation)重要

- 数式, notationにはこだわる, そして, 絶対に間違えないようにする
 - => 数式を間違えている論文は, 内容も間違えている可能性があると思われる

[参考] 数式

書き方に迷う（あるいは自分のスタイルがない）場合はこの法則を真似るとよい

● DLbook notation (ICLR推奨)

https://github.com/goodfeli/dlbook_notation/

https://github.com/ICLR/Master-Template/blob/master/iclr2020/iclr2020_conference.pdf

Numbers and Arrays

a	A scalar (integer or real)
\mathbf{a}	A vector
\mathbf{A}	A matrix
\mathbf{A}	A tensor
\mathbf{I}_n	Identity matrix with n rows and n columns
\mathbf{I}	Identity matrix with dimensionality implied by context
$\mathbf{e}^{(i)}$	Standard basis vector $[0, \dots, 0, 1, 0, \dots, 0]$ with a 1 at position i
$\text{diag}(\mathbf{a})$	A square, diagonal matrix with diagonal entries given by \mathbf{a}
a	A scalar random variable
\mathbf{a}	A vector-valued random variable
\mathbf{A}	A matrix-valued random variable

Sets and Graphs

\mathbb{A}	A set
\mathbb{R}	The set of real numbers
$\{0, 1\}$	The set containing 0 and 1
$\{0, 1, \dots, n\}$	The set of all integers between 0 and n
$[a, b]$	The real interval including a and b
$(a, b]$	The real interval excluding a but including b
$\mathbb{A} \setminus \mathbb{B}$	Set subtraction, i.e., the set containing the elements of \mathbb{A} that are not in \mathbb{B}
\mathcal{G}	A graph
$\text{Pa}_{\mathcal{G}}(x_i)$	The parents of x_i in \mathcal{G}

[参考] 数式

書き方に迷う（あるいは自分のスタイルがない）場合はこの法則を真似るとよい



● DLbook notation (ICLR推奨)

https://github.com/goodfeli/dlbook_notation/

https://github.com/ICLR/Master-Template/blob/master/iclr2020/iclr2020_conference.pdf

Functions

$f : \mathbb{A} \rightarrow \mathbb{B}$	The function f with domain \mathbb{A} and range \mathbb{B}
$f \circ g$	Composition of the functions f and g
$f(\mathbf{x}; \boldsymbol{\theta})$	A function of \mathbf{x} parametrized by $\boldsymbol{\theta}$. (Sometimes we write $f(\mathbf{x})$ and omit the argument $\boldsymbol{\theta}$ to lighten notation)
$\log x$	Natural logarithm of x
$\sigma(x)$	Logistic sigmoid, $\frac{1}{1 + \exp(-x)}$
$\zeta(x)$	Softplus, $\log(1 + \exp(x))$
$\ \mathbf{x}\ _p$	L^p norm of \mathbf{x}
$\ \mathbf{x}\ $	L^2 norm of \mathbf{x}
x^+	Positive part of x , i.e., $\max(0, x)$

Indexing

a_i	Element i of vector \mathbf{a} , with indexing starting at 1
a_{-i}	All elements of vector \mathbf{a} except for element i
$A_{i,j}$	Element i, j of matrix \mathbf{A}
$\mathbf{A}_{i,:}$	Row i of matrix \mathbf{A}
$\mathbf{A}_{:,i}$	Column i of matrix \mathbf{A}
$A_{i,j,k}$	Element (i, j, k) of a 3-D tensor \mathbf{A}
$\mathbf{A}_{:,:,i}$	2-D slice of a 3-D tensor
\mathbf{a}_i	Element i of the random vector \mathbf{a}

参考文献重要

- 参考文献をきちんと入れる。参考文献のセクションを軽視しない
=> 間違えると「わかってない」と思われる
- 関連研究との関係を明確化
 - 研究の立ち位置を明確化
 - 対象/非対象の研究を明確化
- サーベイを (何度でも) きっちりやる

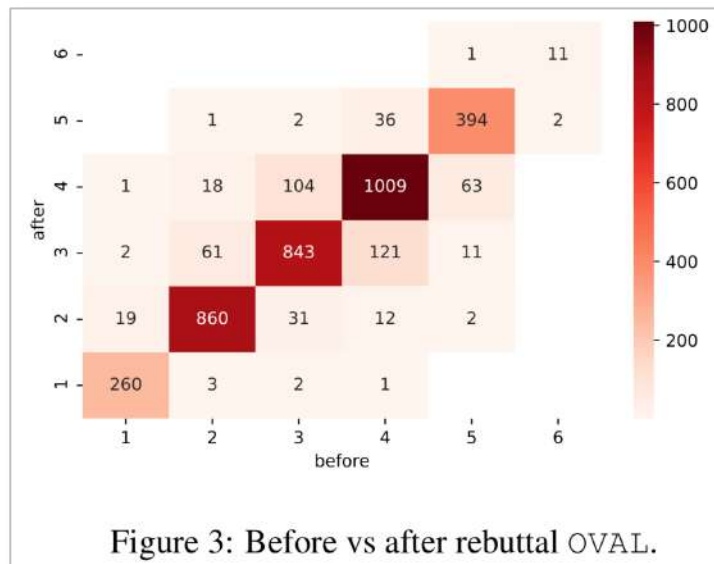
概略図重要

- わかりやすい図
 - 典型的な読者の負荷を下げる方法
- 直感的にわかりやすい図を入れる
 - 特に最初のページに概要図

賢い author response (rebuttal) の取り組み方

Author Response (Rebuttal)

- 基本的にauthor response でスコアを上げてくれる可能性はそれほど高くはない



Does My Rebuttal Matter? Insights from a Major NLP Conference [NAACL-2019]
<https://www.aclweb.org/anthology/N19-1129>

- うまくやればスコア上げてもらえる可能性あり
 - ボーダーライン上の場合には死活問題
- ⇒ ただ聞かれてことを答えるだけではほぼ意味がない

Author Response (Rebuttal) の心得

- 基本姿勢/心得
 - Reviewerは神様です！
 - 論理的かつ**根拠**を示す (感情的になっては絶対ダメ)
 - 根拠：参考文献，統計情報など客観的事実
- 基本戦略
 - **(厚く)回答する質問の狙いを絞る**
=> 全ての質問やnegative reviewに答えるのは難しい
 - 「**あれ，これはもしかして自分が間違えてるかも？**」という気持ちにさせる (ように書く)
=> 基本review内容は変えてくれないと思っておく
 - **議論を促すような書き方をする**

狙いの絞り方

- 手順 (例)

- まず、positive, negative, 意味不明 の要領で review に全てマーカーを入れる
- マーカーの色の分布やreviewのニュアンスから、Reviewerが全体として通してあげようと思っているか、落としにかかっているかを見積もる (overall scoreが出ている場合はそれを見てもよい)
 - 各Reviewer単位の狙い目を見積もるため
- negative review の中で重要度 (致命度?) の順にラ
ンキング
 - Reviewの項目単位での狙い目を見積もるため
- どの negative review に対して反論を書くか決める
- (意味不明にあえてツッコミを入れるかは戦略次第)

狙いの絞り方

- 狙い目の (negative) review
 - 抽象的なもの：根拠が示されていない
 - => 反論に根拠 (参考文献情報や統計など) を示して提示すればよい
- 狙い目のreviewer (?)

前提：全てのnegative reviewに答えるスペースがない

 - 必ずしも最も negative な reviewer に対して重点的に反論するのが正しい戦略ではない
 - Review内容は好意的だが点数にギャップがある reviewerを狙うのも一つのやり方

意見を述べる

- 回答した結果どうなった (と思っている) のかも具体的にかく
 - 「これで全ての疑問点は解決したはず」
- ただし、これはあまりやりすぎると逆効果
 - => 査読者は神様
 - => 査読者コストが上がる

まとめ

まとめ

- 最難関国際会議への投稿/採録に向けた知見共有
 - 査読は非常に高コスト
 - 読み手 (査読者) の負荷が低くなるように心がける
 - 読み手 (査読者) の役に立つ情報となるかを検討する