On the Foundations of Diverse Information Retrieval

Scott Sanner, Kar Wai Lim, Shengbo Guo, Thore Graepel, Sarvnaz Karimi, Sadegh Kharazmi
Outline

• Need for diversity

• The answer: MMR

• But what was the question?
  – Expected n-call@k
Search Result Ranking

We query the daily news for “technology”

Is this desirable?

Note that de-duplication does not solve this problem
Recommendation

• Book search for “cowboys”*

*These are actual results I got from an e-book search engine.

• Why are they mostly romance books?
  – Will this appeal to all demographics?
Diversity Beyond IR: Machine Learning

• Classifying Computer Science web pages
  – Select top features by some feature scoring metric
    • computer
    • computers
    • computing
    • computation
    • computational

• Certainly all are appropriate
  – But do these cover all relevant web pages well?
  – A better approach? MRMR?
Diversity in IR

• In this talk, focus on diversity from an IR perspective:

  – De-duplication (all search engines handle – locality sensitive hashing)
    • Same page, different URL
    • Different page versions (copied Wiki articles)

  – Source diversity (easy)
    • Web pages vs. news vs. image search vs. Youtube

  – Sense ambiguity (easily addressed through user reformulation)
    • Java, Jaguar, Apple
    • Arguably not the main motivation

  – Intrinsic diversity (faceted information needs)
    • Heathrow (checkin, food services, ground transport)

  – Extrinsic diversity (diverse user population)
    • Teens vs. parents, men vs. women, location

How do these relate to previous examples?

Radlinski and Joachims – diverse information needs (SIGIR Forum 2009)
Diversification in IR

- Maximum marginal relevance (MMR)
  - Carbonell & Goldstein, SIGIR 1998
  - *Standard* diversification approach in IR

- MMR Algorithm:
  - $S_k$ is subset of $k$ selected documents from $D$
  - Greedily build $S_k$ from $S_{k-1}$ where $S_0 = \emptyset$ as follows:

$$s_k^* = \arg\max_{s_k \in D \setminus S_{k-1}^*} [\lambda(\text{Sim}_1(q, s_k)) - (1 - \lambda) \max_{s_i \in S_{k-1}^*} \text{Sim}_2(s_i, s_k)]$$
What was the Question?

• MMR is an algorithm, we don’t know what underlying objective it is optimizing.

• Previous formalization attempts but full question unanswered for 14 years
  – Chen and Karger, SIGIR 2006 came closest

• This talk: a complete derivation of MMR
  – Many assumptions
  – Arguably the assumptions you are making when using MMR!
Where do we start?

Let’s try to relate set/ranking objective Precision@k to diversity*

*Note: non-standard IR! IR evaluates these objectives empirically but never derives algorithms to directly optimize them! (Largely because long tail queries & no labels.)
Relating Precision@k Objectives to Diversity

• Chen and Karger, SIGIR 2006: **1-call@k**
  – At least one document in $S_k$ should be relevant ($P@k=1$)
  – **Very Diverse**: encourages you to “cover your bases” with $S_k$
    • *Sanner et al*, CIKM 2011: 1-call@k derives MMR with $\lambda = ½$

• van Rijsbergen, 1979: **Probability Ranking Principle (PRP)**
  – Rank items by probability of relevance (e.g., modeled via term freq)
    • PRP relates to k-call@k ($P@k=k$) which relates to MMR with $\lambda = 1$
  – **Not diverse**: Encourages $k^{th}$ item to be very similar to first $k-1$ items

• So either $\lambda = ½$ (1-call@k – very diverse) or $\lambda = 1$ (k-call@k – not diverse)?
  – Should really tune $\lambda$ for MMR based on query ambiguity
    • *Santos, MacDonald, Ounis*, CIKM 2011: Learn best $\lambda$ given query features
  – So what derives $\lambda \in [½, 1]$?
    • Any guesses? 😊

Small fraction of queries have diverse information needs – need good experimental design
Empirical Study of n-call@k

• How does diversity of n-call@k change with n?

Clearly, diversity (pairwise document correlation) decreases with n in n-call@k

J. Wang and J. Zhu. Portfolio theory of information retrieval, SIGIR 2009
Hypothesis

• Let’s try optimizing 2-call@k
  – Derivation builds on Sanner et al, CIKM 2011
  – Optimizing this leads to MMR with \( \lambda = \frac{2}{3} \)

• There seems to be a trend relating \( \lambda \) and \( n \):
  – \( n=1: \lambda = \frac{1}{2} \)
  – \( n=2: \lambda = \frac{2}{3} \)
  – \( n=k: 1 \)

• Hypothesis
  – Optimizing n-call@k leads to MMR with \( \lim_{k \to \infty} \lambda(k,n) = \frac{n}{n+1} \)
Recap

• We wanted to know what objective leads to MMR diversification

• Evidence supports that optimizing $n$-call@$k$ leads to diverse MMR-like behavior where $\lambda = \frac{n}{n+1}$

• Can we derive MMR from $n$-call@$k$?
One Detail is Missing...

• We want to optimize n-call@k
  – i.e., at least n of k documents should be relevant
  – Great, but given a query and corpus, how do we do this?

• Key question: how to define “relevance”?
  – Need a model for this – probabilistic given PRP connections
  – If diversity needed to cover latent information needs

→ relevance model must include latent query/doc “topics”
Latent Binary Relevance Retrieval Model

• Formulate as optimization in the graphical model:
  – $s_i$: doc selection $i=1..k$
  – $t_i$: topic for $i$
  – $r_i$: $i$ relevant?
  – $q$: query
  – $t'$: topic for $q$
How to determine latent topics?

- **observed**
- **latent**

- Need CPTs for
  - \( P(t_i \mid s_i) \)
  - \( P(t' \mid q) \)

- Can...
  - Set arbitrarily
    - Topics are words
    - L1-norm TF or TF-IDF!
  - Topic modeling (not quite LDA)
Defining Relevance

- Adapt 0-1 loss model of PRP:

\[
P(r_i|t', t_i) = \begin{cases} 
0 & \text{if } t_i \neq t' \\
1 & \text{if } t_i = t'
\end{cases}
\]
Optimizing Expected 1-call@k

\[ S^* = \arg\max_{S=\{s_1, \ldots, s_k\}} \text{Exp-1-call@k}(S, \vec{q}) \]

\[
\text{Exp-1-call@k}(S, \vec{q}) \\
= \mathbb{E}\left[ \bigvee_{i=1}^{k} r_i = 1 \middle| s_1, \ldots, s_k, \vec{q} \right] = P(\bigvee_{i=1}^{k} r_i = 1|s_1, \ldots, s_k, \vec{q}) \\
= P(\left[ r_1 = 1 \right] \vee \left[ r_1 = 0 \land r_2 = 1 \right] \vee \left[ r_1 = 0 \land r_2 = 0 \land r_3 = 1 \right] \vee \ldots |s_1, \ldots, s_k, \vec{q}) \\
= \sum_{i=1}^{k} P(r_i = 1, r_1 = 0, \ldots, r_{i-1} = 0|s_1, \ldots, s_k, \vec{q}) \\
= \sum_{i=1}^{k} P(r_i = 1|r_1 = 0, \ldots, r_{i-1} = 0, s_1, \ldots, s_k, \vec{q})P(r_1 = 0, \ldots, r_{i-1} = 0|S, \vec{q})
\]

Greedy: \( s_i^* = \arg\max_{s_i} P(r_i = 1|r_1 = 0, \ldots, r_{i-1} = 0, s_1^*, \ldots, s_{i-1}^*, s_i, \vec{q}) \)

All disjuncts mutually exclusive

s_k D-separated from \(r_1 \ldots r_{k-1}\); so can ignore when greedy!
Objective to Optimize: \( s_1^* \)

- Take a greedy approach (like MMR)
- Choose \( s_1 \) via AccRel first

\[
s_1^* = \arg \max_{s_1} P(r_1|s_1, \tilde{q})
\]

\[
= \arg \max_{s_1} \sum_{t_1, t'} \mathbb{1}[t' = t_1] P(t'|\tilde{q}) P(t_1|s_1)
\]

\[
= \arg \max_{s_1} \sum_{t'} P(t'|\tilde{q}) P(t_1 = t'|s_1)
\]

Can derive numerous kernels including TF, TD-IDF, LSI
Objective to Optimize: $S_2^*$

- Choose $s_2$ via AccRel next
  - Condition on chosen $s_1^*$ and $r_1=0$

\[
S^*_2 = \arg \max_{s_2} P(r_2 = 1| r_1 = 0, s_1^*, s_2, \bar{q})
\]

\[
= \arg \max_{s_2} \sum_{t_1, t_2, t'} \mathbb{I}[t_2 = t']P(t_1|s_1^*)\mathbb{I}[t_1 \neq t']P(t_2|s_2)P(t'|\bar{q})
\]

\[
= \arg \max_{s_2} \sum_{t'} p(t'|\bar{q})P(t_2 = t'|s_2)(1 - P(t_1 = t'|s_1^*))
\]

\[
= \arg \max_{s_2} \left[ \sum_{t'} p(t'|\bar{q})P(t_2 = t'|s_2) \right] - \left[ \sum_{t'} P(t'|\bar{q})p(t_1 = t'|s_1^*)P(t_2 = t'|s_2) \right]
\]

What is $\lambda$? $\frac{1}{2}$. 

Query-topic weighted diversity!
Objective to Optimize: $S_k^*$, $k>2$

$$s_k^* = \arg \max_{s_k \in D \setminus S_{k-1}} \sum_{t'} P(t_k = t'|s_k)P(t'|q) \prod_{i=1}^{k-1} [1 - P(t_i = t'|s_i^*)]$$

$$\prod_{i=1}^{k-1} [1 - P(t_i = t'|s_i^*)] = 1 - \left[ \sum_{i=1}^{k-1} P(t_i = t'|s_i^*) - \sum_{i=1}^{k-1} \sum_{j=1, j \neq i}^{k-1} P(t_i = t'|s_i^*)P(t_j = t'|s_j^*) + \ldots \right]$$

Derives topic $t'$ coverage by Principle of Inclusion, Exclusion!
Provides set-covering view of diversity.
So far...

• We’ve seen hints of MMR from $E[1\text{-call}@k]$
  – Need a few more assumptions to get to MMR

• Let’s also generalize to $E[n\text{-call}@k]$ for general $\lambda$:

$$\text{Exp-}n\text{-Call}@k(S_k, q) = \mathbb{E}[R_k \geq n | s_1, \ldots, s_k, q]$$

where

$$R_k = \sum_{i=1}^k r_i$$
Optimization Objective

• Continue with greedy approach for $E[n \cdot \text{call} @ k]$

  – Select the next document $s_k^*$ given all previously chosen documents $S_{k-1}$:

    $$s_k^* = \arg \max_{s_k} \mathbb{E}[R_k \geq n | S_{k-1}^*, s_k, q]$$
Derivation

• Nontrivial
  – Only an overview of “key tricks” here

• For full details, see
  – Sanner et al, CIKM 2011: 1-call@k (gentler introduction)
  – Lim et al, SIGIR 2012: n-call@k
  and online SIGIR 2012 appendix
$$s_k^* = \arg \max_{s_k} \mathbb{E}[R_k \geq n | S_{k-1}^*, s_k, q]$$

$$= \arg \max_{s_k} P(R_k \geq n | S_{k-1}^*, s_k, q)$$
Derivation

\[ s^*_k = \arg \max_{s_k} \mathbb{E}[R_k \geq n | S^*_{k-1}, s_k, q] \]

\[ = \arg \max_{s_k} P(R_k \geq n | S^*_{k-1}, s_k, q) \]

\[ = \arg \max_{s_k} \sum_{T_k} \left( P(t_k|q) \cdot P(t_k|s_k) \prod_{i=1}^{k-1} P(t_i|s^*_i) \cdot P(R_k \geq n | T_k, S^*_{k-1}, s_k, q) \right) \]

Marginalise out all subtopics (using conditional probability)

\[ T_k = \{t, t_1, \ldots, t_k\} \text{ and } \sum_{T_k} = \sum_t \sum_{t_1} \cdots \sum_{t_k} \]
We write $r_k$ as conditioned on $R_{k-1}$, where it decomposes into two independent events, hence the +
Derivation

\[ s_k^* = \arg \max_{s_k} \mathbb{E}[R_k \geq n|S_{k-1}^*, s_k, q] \]

\[ = \arg \max_{s_k} P(R_k \geq n|S_{k-1}^*, s_k, q) \]

\[ = \arg \max_{s_k} \sum_{T_k} \left( P(t|q) P(t_k|s_k) \prod_{i=1}^{k-1} P(t_i|s_i^*) \cdot P(R_k \geq n|T_k, S_{k-1}^*, s_k, q) \right) \]

\[ = \arg \max_{s_k} \sum_{T_k} P(t|q) P(t_k|s_k) \prod_{i=1}^{k-1} P(t_i|s_i^*) \]

\[ \cdot \left( P(r_k = 0|R_{k-1} \geq n, t_k, t) P(R_{k-1} \geq n|T_{k-1}) \right. \]

\[ + P(r_k = 1|R_{k-1} = n-1, t_k, t) P(R_{k-1} = n-1|T_{k-1}) \bigg) \]

\[ = \arg \max_{s_k} \left( \sum_{T_{k-1}} \left[ \sum_{t_k} \right. P(t_k|s_k) P(R_{k-1} \geq n|T_{k-1}) P(t|q) \prod_{i=1}^{k-1} P(t_i|s_i^*) + \right. \]

\[ \left. \sum_{t} P(t|q) P(t_k = t|s_k) \sum_{t_{1,\ldots,t_{k-1}}} P(R_{k-1} = n-1|T_{k-1}) \prod_{i=1}^{k-1} P(t_i|s_i^*) \right) \]

\[ \sum_{t_k} P(t_k|s_k) P(r_k=1|t_k, t) \]

\[ = \sum_{t_k} P(t_k|s_k) \mathbb{I}[t_k=t] = P(t_k=t|s_k) \]

Start to push latent topic marginalizations as far in as possible.
Derivation

$$s_k^* = \arg \max_{s_k} \mathbb{E}[R_k \geq n \mid S_{k-1}^*, s_k, q]$$

$$= \arg \max_{s_k} P(R_k \geq n \mid S_{k-1}^*, s_k, q)$$

$$= \arg \max_{s_k} \sum_{T_k} \left( P(t \mid q) P(t_k \mid s_k) \prod_{i=1}^{k-1} P(t_i \mid s_i^*) \cdot P(R_k \geq n \mid T_k, S_{k-1}^*, s_k, q) \right)$$

$$= \arg \max_{s_k} \sum_{T_k} P(t \mid q) P(t_k \mid s_k) \prod_{i=1}^{k-1} P(t_i \mid s_i^*) \cdot \left( P(r_k \geq 0 \mid R_{k-1} \geq n, t_k, t) P(R_{k-1} \geq n \mid T_{k-1}) \right. \right.$$

$$+ P(r_k = 1 \mid R_{k-1} = n-1, t_k, t) P(R_{k-1} = n-1 \mid T_{k-1}) \right)$$

$$= \arg \max_{s_k} \left( \sum_{T_{k-1}} \left[ \sum_{t_k} P(t_k \mid s_k) \right] P(R_{k-1} \geq n \mid T_{k-1}) P(t \mid q) \prod_{i=1}^{k-1} P(t_i \mid s_i^*) \right. \right.$$  

$$+ \sum_{t} P(t \mid q) P(t_k = t \mid s_k) \sum_{t_1, \ldots, t_{k-1}} P(R_{k-1} = n-1 \mid T_{k-1}) \prod_{i=1}^{k-1} P(t_i \mid s_i^*) \right)$$

$$= \arg \max_{s_k} \sum_{t} P(t \mid q) P(t_k = t \mid s_k) P(R_{k-1} = n-1 \mid S_{k-1}^*)$$

First term in + is independent of $s_k$ so can remove from max!
Derivation

• We arrive at the simplified

\[ s_k^* = \arg \max_{s_k} \mathbb{E}[R_k \geq n | S_{k-1}^*, s_k, q] \]

\[ = \arg \max_{s_k} \sum_t P(t|q)P(t_k = t|s_k)P(R_{k-1} = n-1 | S_{k-1}^*) \]

• This is still a complicated expression, but it can be expressed recursively...
Recursion

\[ P(R_k = n | S_k, t) = \]
\[
\begin{cases}
  n \geq 1, k > 1 : & (1 - P(t_k = t | s_k)) P(R_{k-1} = n | S_{k-1}, t) \\
  & + P(t_k = t | s_k) P(R_{k-1} = n - 1 | S_{k-1}, t) \\
  n = 0, k > 1 : & (1 - P(t_k = t | s_k)) P(R_{k-1} = 0 | S_{k-1}, t) \\
  n = 1, k = 1 : & P(t_1 = t | s_1) \\
  n = 0, k = 1 : & 1 - P(t_1 = t | s_1) \\
  n > k : & 0
\end{cases}
\]

Very similar conditional decomposition as done in first part of derivation.
Unrolling the Recursion

• We can unroll the previous recursion, express it in closed-form, and substitute:

\[
 s_k^* = \arg \max_{s_k} \sum_t \left( P(t|q) P(t_k = t|s_k) \sum_{j_1, \ldots, j_{n-1}} \prod_{l \in \{j_1, \ldots, j_{n-1}\}} P(t_l = t|s_l^*) \prod_{i=1, i \not\in \{j_1, \ldots, j_{n-1}\}} P(t_i = t|s_i^*) \right)
\]

Where's the max? MMR has a max.

\[
 n \leq k/2
\]

\[
 s_k^* = \arg \max_{s_k} \sum_t \left( P(t|q) P(t_k = t|s_k) \sum_{j_n, \ldots, j_{k-1}} \prod_{l \in \{j_n, \ldots, j_{k-1}\}} (1 - P(t_l = t|s_l^*)) \prod_{i=1, i \not\in \{j_n, \ldots, j_{k-1}\}} P(t_i = t|s_i^*) \right)
\]

where \( j_1, \ldots, j_{n-1} \in \{1, \ldots, k - 1\} \) satisfy that \( j_i < j_{i+1} \)
Deterministic Topic Probabilities

• We assume that the topics of each document are known (deterministic), hence:

\[ P(t_i | s_i) \in \{0, 1\} \]

– Likewise for \( P(t|q) \)
– This means that a document refers to exactly one topic and likewise for queries, e.g.,
  • If you search for “Apple” you meant the fruit OR the company, but not both
  • If a document refers to “Apple” the fruit, it does not discuss the company Apple Computer
Deterministic Topic Probabilities

• Generally:

\[
\begin{bmatrix}
P(t_i = C_1 | s_i) \\
P(t_i = C_2 | s_i) \\
\vdots \\
P(t_i = C_{|T|} | s_i)
\end{bmatrix} = \begin{bmatrix}
0.24 \\
0.62 \\
\vdots \\
0.01
\end{bmatrix}
\]

• Deterministic:

\[
\begin{bmatrix}
P(t_i = C_1 | s_i) \\
P(t_i = C_2 | s_i) \\
\vdots \\
P(t_i = C_{|T|} | s_i)
\end{bmatrix} = \begin{bmatrix}
0 \\
1 \\
\vdots \\
0
\end{bmatrix}
\]
Convert a $\Pi$ to a max

- Assuming deterministic topic probabilities, we can convert a $\Pi$ to a max and vice versa

- For $x_i \in \{0 \text{ (false)}, 1 \text{ (true)}\}$
  
  $\max_i = \vee_i x_i$
  
  $= \neg \land_i (\neg x_i)$
  
  $= 1 - \land_i (1 - x_i)$
  
  $= 1 - \prod_i (1 - x_i)$
Convert a $\prod$ to a max

- From the optimizing objective when $n \leq k/2$, we can write

$$\prod_{\substack{i=1 \\ i \notin \{j_1, \ldots, j_{n-1}\}}}(1 - P(t_i = t|s_i^*)) = 1 - \left(1 - \prod_{\substack{i=1 \\ i \notin \{j_1, \ldots, j_{n-1}\}}}(1 - P(t_i = t|s_i^*))\right)$$

$$= 1 - \left(\max_{\substack{i \in [1, k-1] \\ i \notin \{j_1, \ldots, j_{n-1}\}}} P(t_i = t|s_i^*)\right)$$
Objective After $\prod \rightarrow \max$

$$s_k^* = \arg\max_{s_k} \sum_t \left( P(t|q) P(t_k = t|s_k) \sum_{j_1,\ldots,j_{n-1}} \prod_{l \in \{j_1,\ldots,j_{n-1}\}} P(t_l = t|s_l^*) \prod_{i=1 \atop i \notin \{j_1,\ldots,j_{n-1}\}}^{k-1} (1 - P(t_i = t|s_i^*)) \right)$$

$$= \arg\max_{s_k} \sum_t \left( P(t|q) P(t_k = t|s_k) \sum_{j_1,\ldots,j_{n-1}} \prod_{l \in \{j_1,\ldots,j_{n-1}\}} P(t_l = t|s_l^*) \right. - P(t|q) P(t_k = t|s_k) \sum_{j_1,\ldots,j_{n-1}} \prod_{l \in \{j_1,\ldots,j_{n-1}\}} P(t_l = t|s_l^*) \left. \max_{i \in [1,k-1]} P(t_i = t|s_i^*) \right) \prod_{i \notin \{j_1,\ldots,j_{n-1}\}}^{k-1} (1 - P(t_i = t|s_i^*)) \right)$$
Combinatorial Simplification

- Deterministic topics also permit combinatorial simplification of some of the $\prod$

- Assuming that $m$ documents out of the chosen $(k-1)$ are relevant, then

\[
\sum_{j_1,\ldots,j_{n-1}} \prod_{l \in \{j_1,\ldots,j_{n-1}\}} P(t_l = t | s^*_l) \text{ (the top term) are non-zero } \binom{m}{n-1} \text{ times.}
\]

- \[
\sum_{j_1,\ldots,j_{n-1}} \prod_{l \in \{j_1,\ldots,j_{n-1}\}} P(t_l = t | s^*_l) \max_{i \in [1,k-1]} P(t_i = t | s^*_i) \text{ (bottom term) are non-zero } \binom{m}{n} \text{ times.}
\]
After...

- assuming a deterministic topic distribution,
- converting $\Pi$ to a max, and
- combinatorial simplification

$$
\text{arg max}_{s_k} \left( \frac{m}{n-1} \right) \sum_t P(t|q)P(t_k=t|s_k) - \left( \frac{m}{n} \right) \max_{s_i \in S^*_k-1} \sum_t P(t_i=t|s_i)P(t|q)P(t_k=t|s_k)
$$

relevance: $\text{Sim}_1(s_k,q)$

$$
\text{arg max}_{s_k} \frac{n}{m+1} \text{Sim}_1(s_k,q) - \frac{m-n+1}{m+1} \max_{s_i \in S^*_k-1} \text{Sim}_2(s_k,s_i,q)
$$

diversity: $\text{Sim}_2(s_k,s_i,q)$

Topic marginalization leads to probability product kernel $\text{Sim}_1(\cdot, \cdot)$: this is any kernel that $L_1$ normalizes inputs, so can use with TF, TF-IDF! MMR drops $q$ dependence in $\text{Sim}_2(\cdot, \cdot)$.

argmax invariant to constant multiplier, use Pascal’s rule to normalize coefficients to $[0,1]$:

$$
\binom{m}{n-1} + \binom{m}{n} = \binom{m+1}{n}
$$
Comparison to MMR

• The optimising objective used in MMR is

\[ s_k^* = \arg \max_{s_k \in D \setminus S_{k-1}^*} \left[ \lambda (\text{Sim}_1(q, s_k)) - (1 - \lambda) \max_{s_i \in S_{k-1}^*} \text{Sim}_2(s_i, s_k) \right] \]

• We note that the optimising objective for expected n-call@k has the same form as MMR, with \( \lambda = \frac{n}{m+1} \).
  – but \( m \) is unknown
Expectation of m

• m is expected number of relevant documents \((m \geq n)\), we can lower bound m as \(m \approx n\).

• With the assumption \(m=n\), we obtain \(\lambda = \frac{n}{n+1}\)
  – Our hypothesis!

\[\lambda = \frac{n}{n+1}\] also roughly follows empirical behavior observed earlier, variation is likely due to m for each corpus

– If instead m constant, still yields MMR-like algorithm
Summary of Contributions

• We derived MMR from n-call@k!

  – After 14 years, we have insight as to what MMR is optimizing!

  – Don’t like the assumptions?
    • Write down the objective you want
    • Derive the solution!
Bigger Picture: Prob ML for IR

• Search engines are complex beasts
  – Manually optimized
    (which has grown out of empirical IR philosophy)

• But there are probabilistic derivations for popular algorithms in IR
  – TF-IDF, BM25, Language Model

• Opportunity for more modeling, learning, optimization
  – Probabilistic models of (latent) information needs
  – And solutions which autonomously learn and optimize these needs!