Topic Models with Logical Constraints on Words

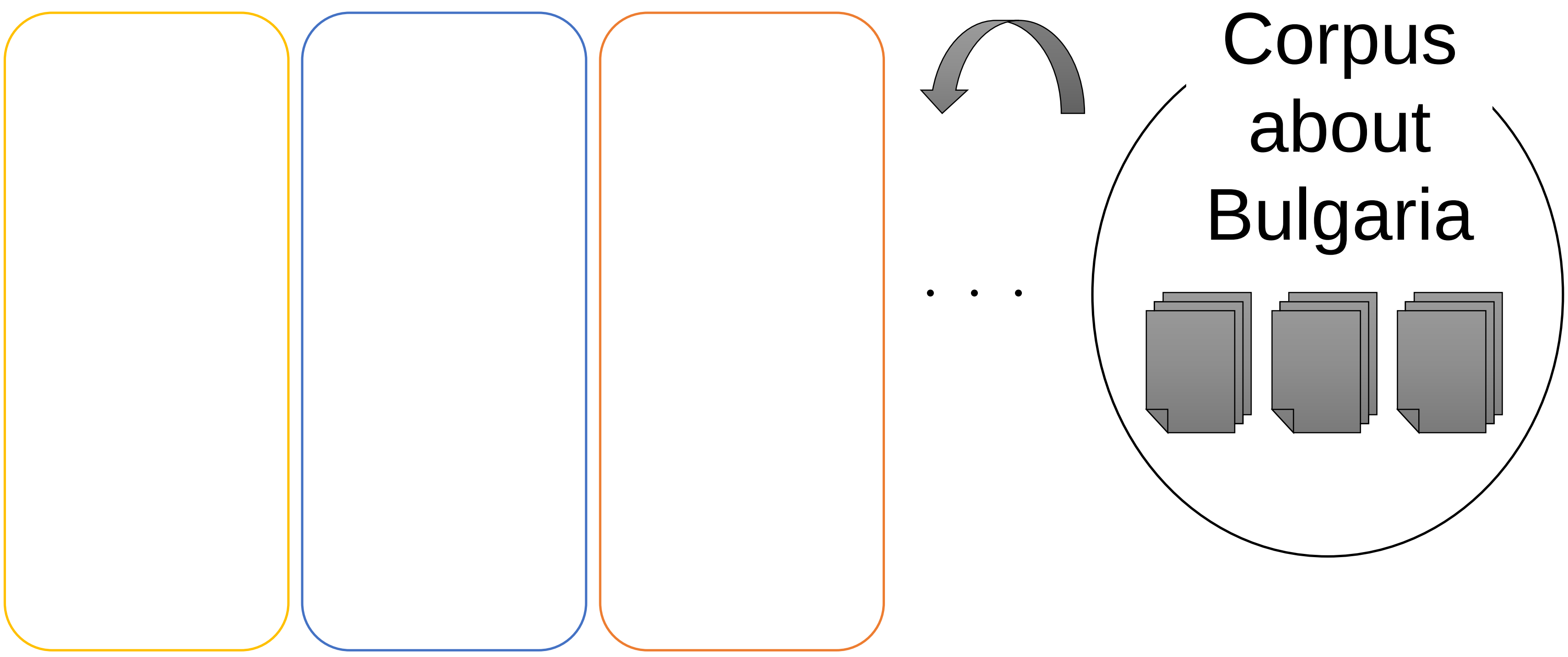
Hayato Kobayashi, Hiromi Wakaki, Tomohiro Yamasaki, and Masaru Suzuki

Corporate Research and Development Center,

Toshiba Corporation, Japan

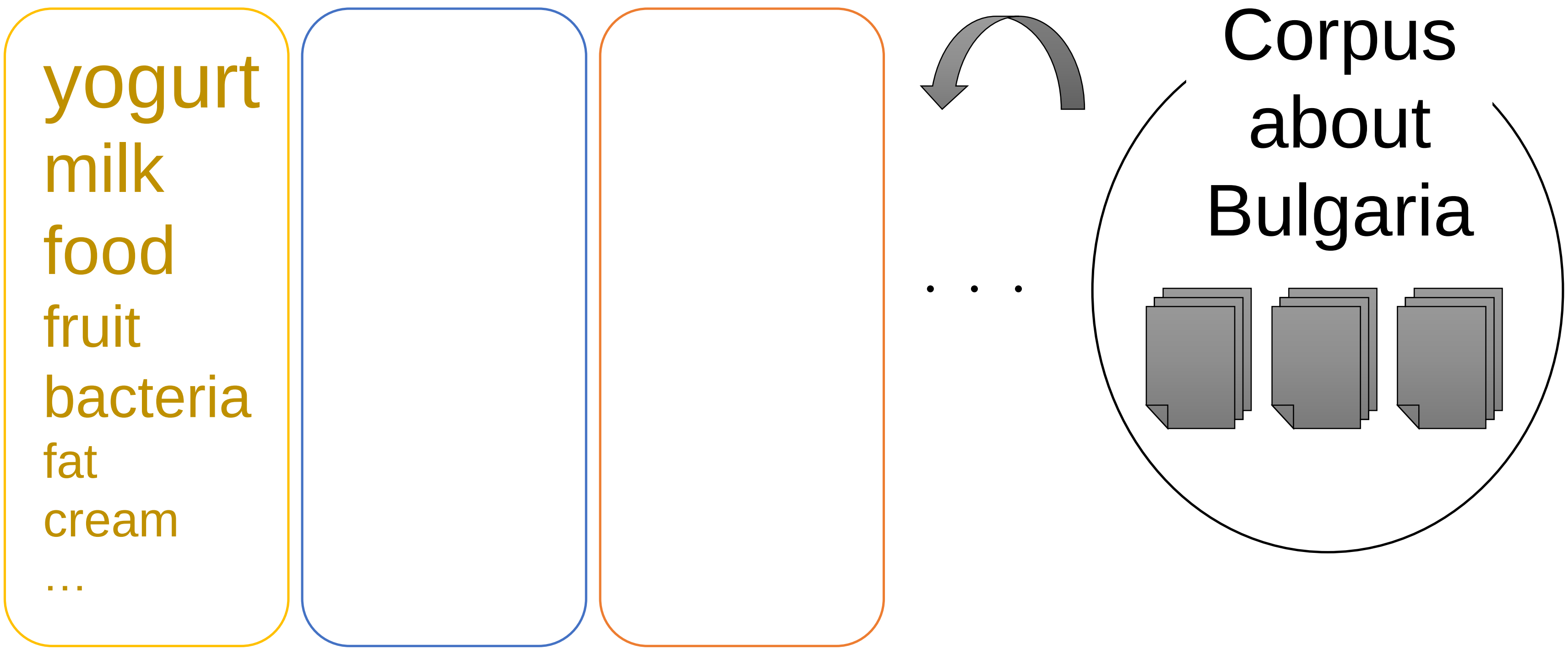
Topic modeling = Word clustering

- Method to extract latent topics on a corpus
  - Each topic is a distribution on words



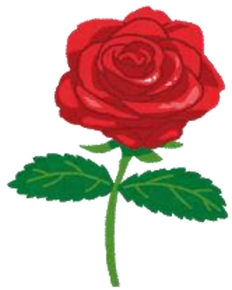
Topic modeling = Word clustering

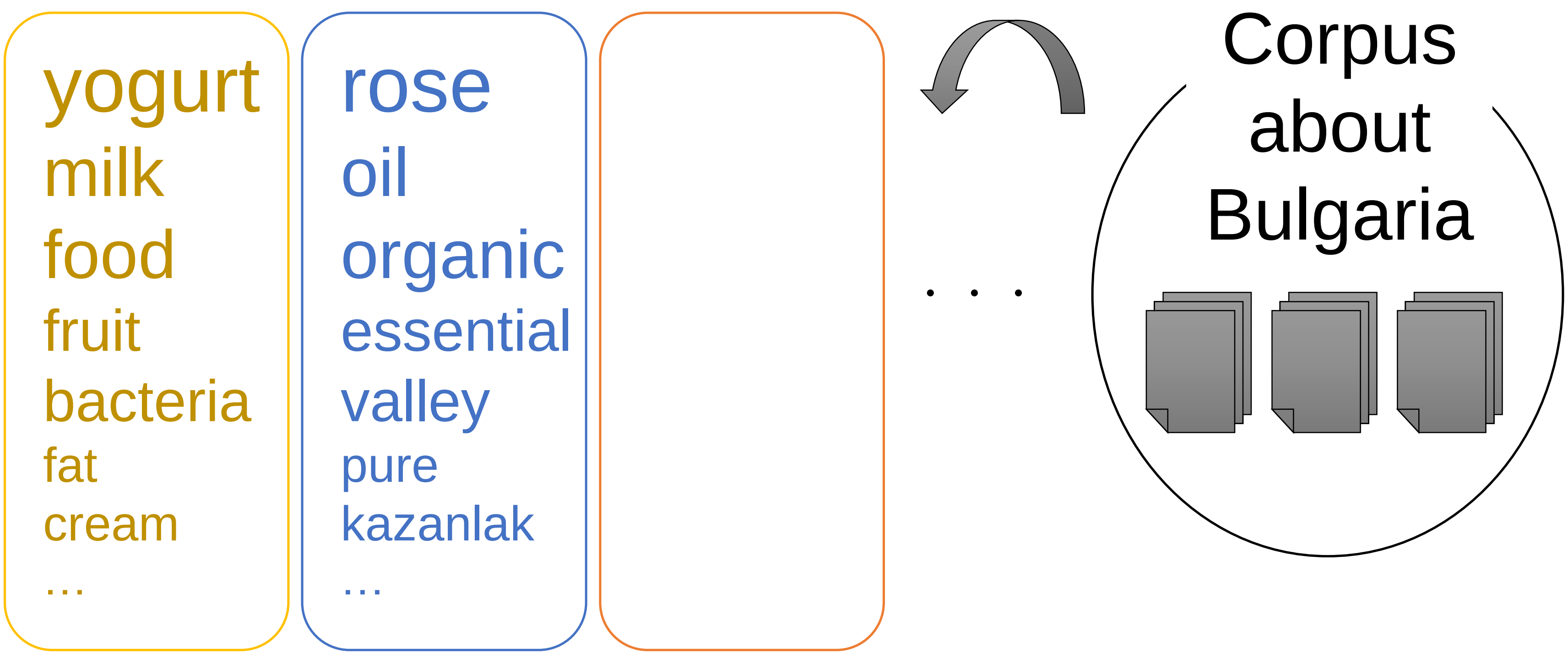
- Method to extract latent topics on a corpus
  - Each topic is a distribution on words



Topic modeling = Word clustering

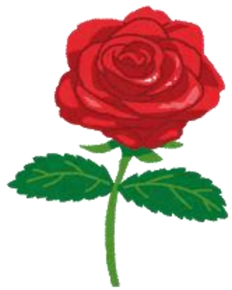
- Method to extract latent topics on a corpus
  - Each topic is a distribution on words

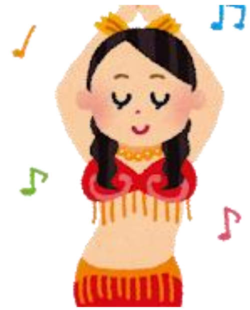




Topic modeling = Word clustering

- Method to extract latent topics on a corpus
  - Each topic is a distribution on words





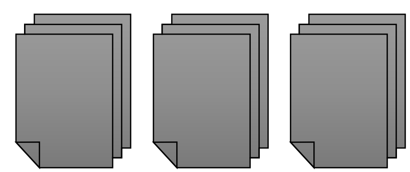
yogurt
milk
food
fruit
bacteria
fat
cream
…

rose
oil
organic
essential
valley
pure
kazanlak
…

dance
fire
sexy
ancient
bikini
walk
exotic
…

. . .

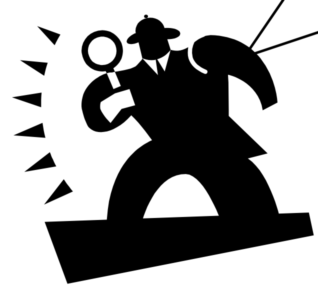
Corpus about Bulgaria



Size of each word represents its frequency

dance
fire
sexy
ancient
bikini
walk
exotic
…

Want to split into “fire dance” and “sexy dance”



Existing work [Andrzejewski+ ICML2009]

- Constraints on words for topic modeling
  - Must-Link(A,B)： A and B appear in the same topic
  - Cannot-Link(A,B)： A and B don’t appear in the same topic

dance		dance
fire	CL	sexy
ancient		bikini
walk		exotic
…		…

Want to split into “fire dance” and “sexy dance”

Cannot-Link(fire, sexy)

Problem of the existing work

- Constraints often don't align with user's intention

You might get "blaze" topic
instead of "fire dance" topic

blaze
fire
ancient
forest
...

CL

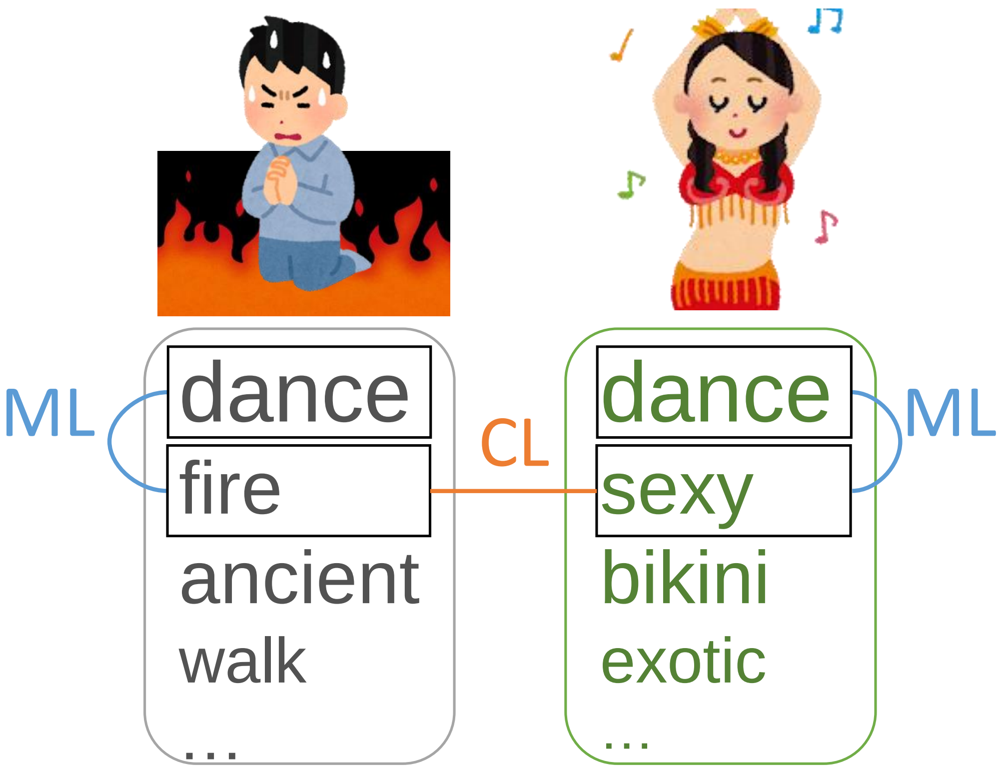
dance
sexy
bikini
exotic
...

Want to split into "fire dance"
and "sexy dance"

Cannot-Link(fire, sexy)

This work

- <u>Logical</u> constraints on words for topic modeling
  - Conjunctions (∧), disjunctions (∨), negations (¬)



Want to split into "fire dance" and "sexy dance"

Cannot-Link(fire, sexy)
∧(Must-Link(dance, fire)
∨ Must-Link(dance, sexy))

Outline of the rest of this talk

- LDA [Blei+ JMLR2003]
  - One of topic modeling method
- LDA-DF [Andrzejewski+ ICML2009]
  - Must-Link and Cannot-Link
- This work
  - Logical expressions of Must-Links and Cannot-Links
  - Experiment
- Conclusion

Latent Dirichlet Allocation (LDA) [Blei+ JMLR2003]

- Famous Topic modeling method

(i) Assume a generative model of documents

- Each topic is a distribution on words
- Each document is a distribution on topics
  - Taken from Dirichlet distributions to generate discrete distributions

(ii) Infer parameters for the two distributions inverting the generative model

Generative process of documents in LDA

- Each topic is a distribution on words
- Each document is a distribution on topics

Topic 1

Document 1

Topic 2

Document 2

Generative process of documents in LDA

- Each topic is a distribution on words
- Each document is a distribution on topics

Topic 1

yogurt
milk
food
fruit
...

Topic 2

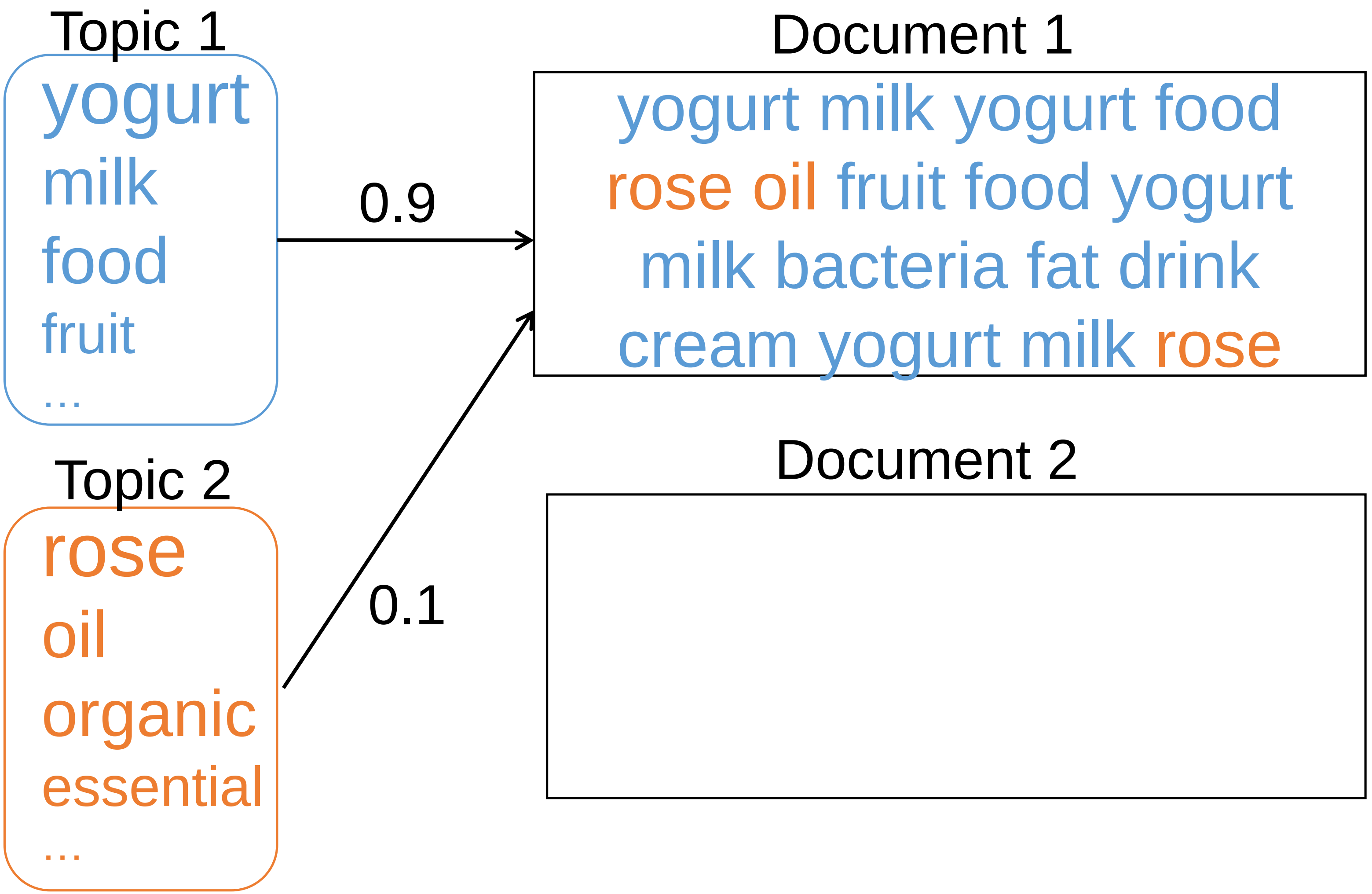
rose
oil
organic
essential
...

Document 1

Document 2

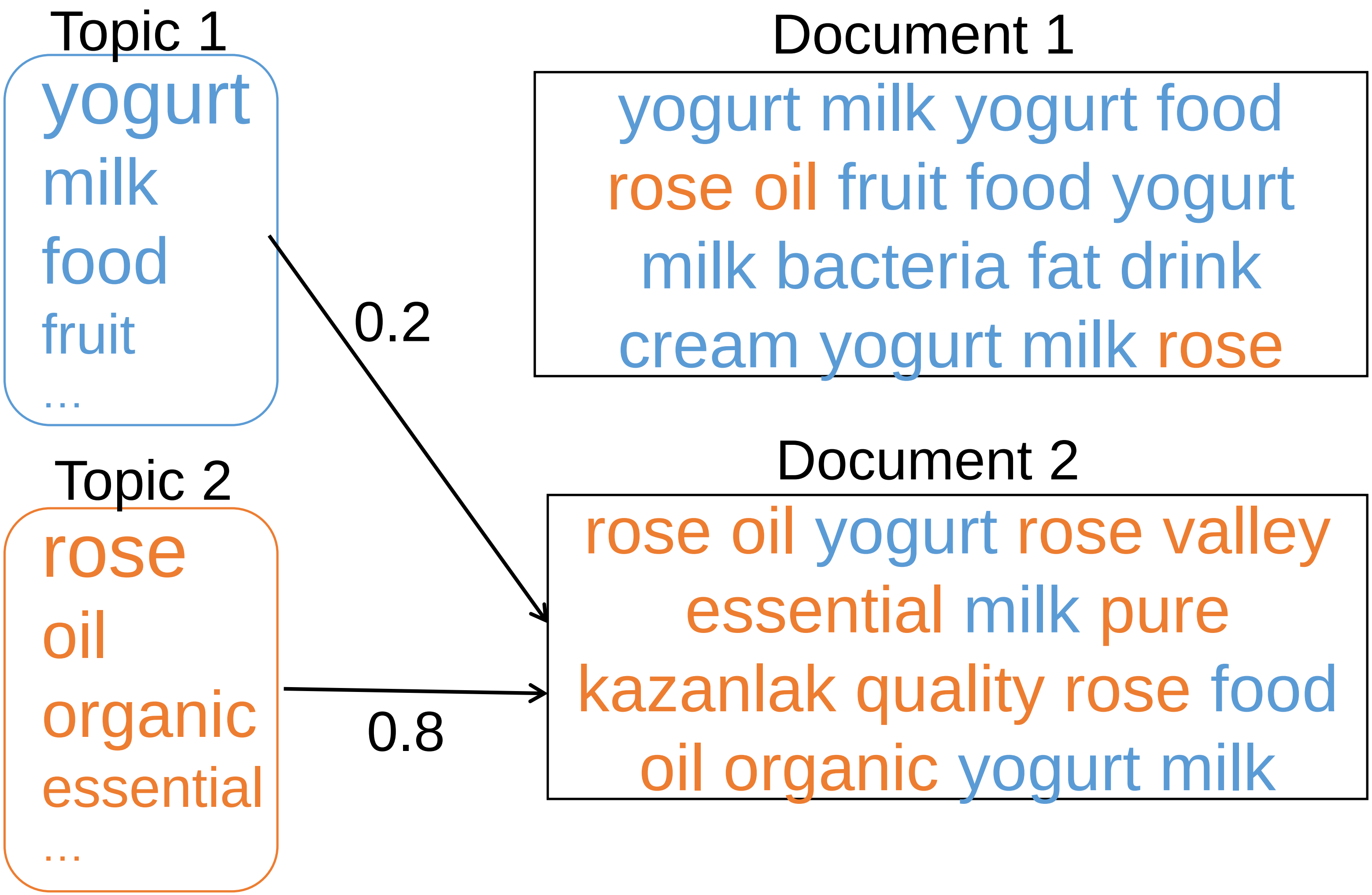
Generative process of documents in LDA

- Each topic is a distribution on words
- Each document is a distribution on topics



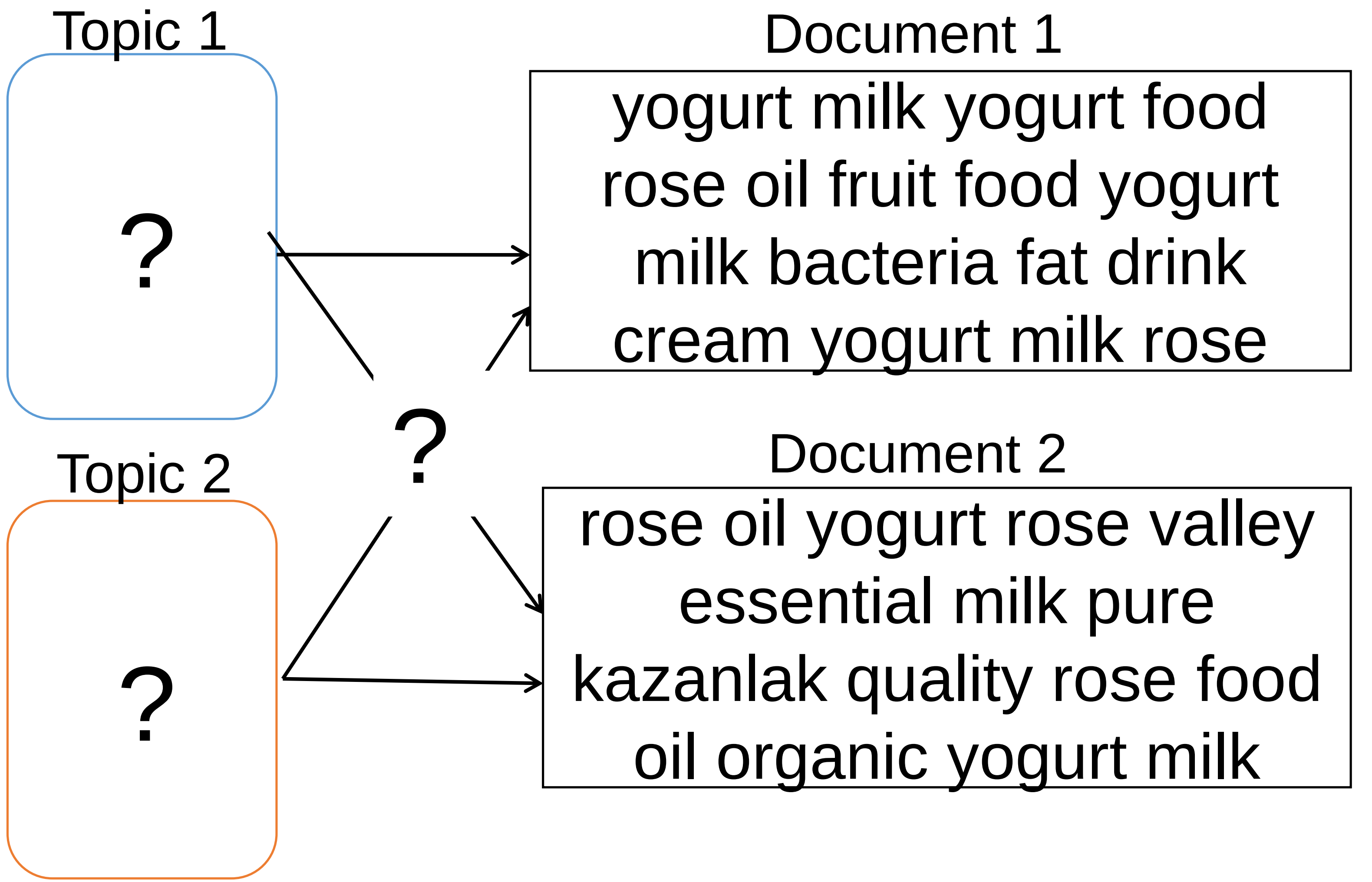
Generative process of documents in LDA

- Each topic is a distribution on words
- Each document is a distribution on topics



Parameter inference in LDA

- Infer word and topic distributions from a corpus inverting the generative process



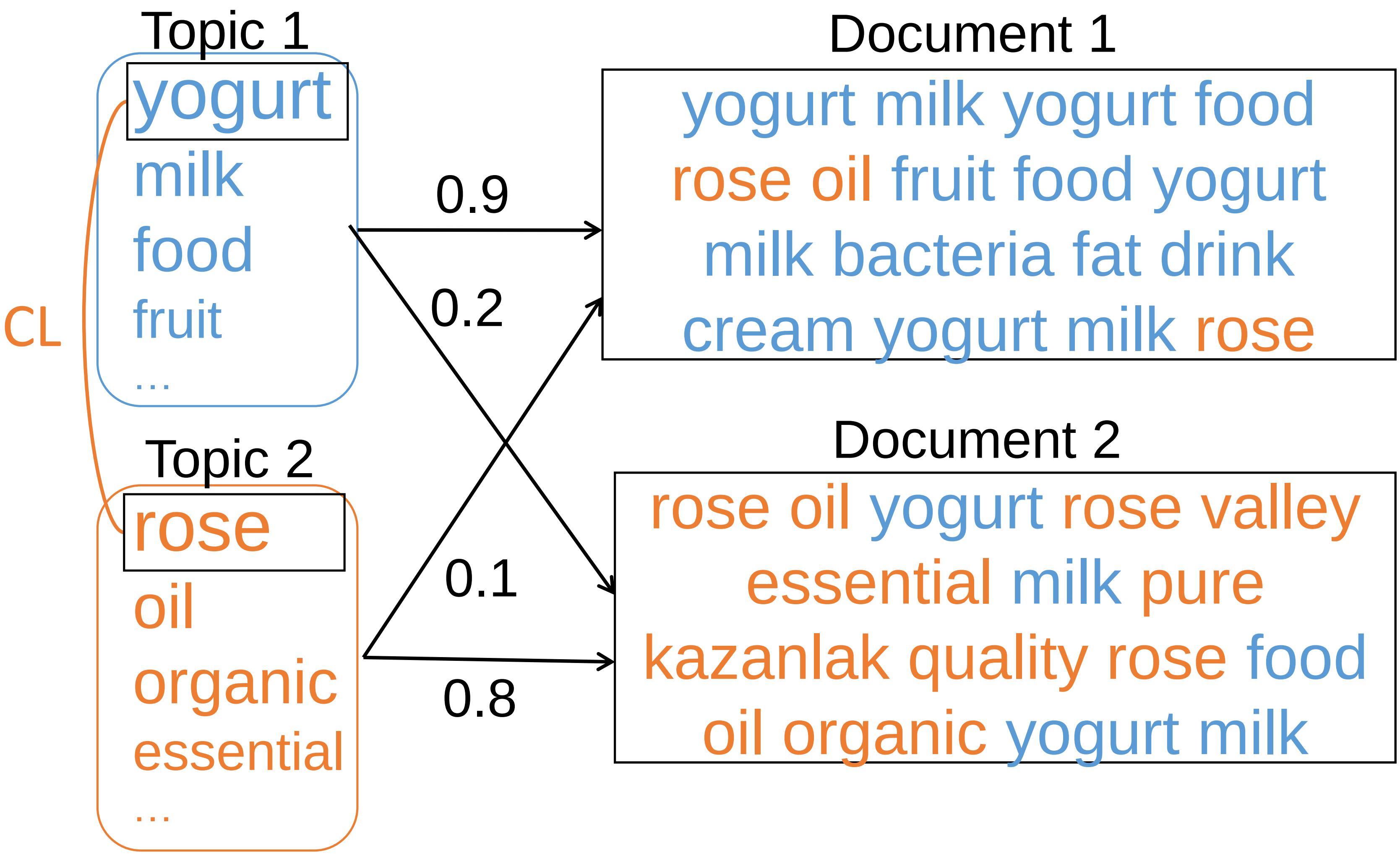
LDA-DF [Andrzejewski+ ICML2009]

- Semi-supervised extension of LDA
  - Only conjunction of Must-Links and Cannot-Links
    - Must-Link(A,B)： A and B appear in the same topic
    - Cannot-Link(A,B)： A and B don't appear in the same topic

- Extending the generative process
  - Each topic is a **constrained** distribution on words
    - Taken from a Dirichlet tree distribution, which is a generalization of a Dirichlet distribution
  - Each document is a distribution on topics
    - Taken from a Dirichlet distribution

Generative process of LDA-DF

- Always generates a distribution, where yogurt and rose do not appear in the same topic.



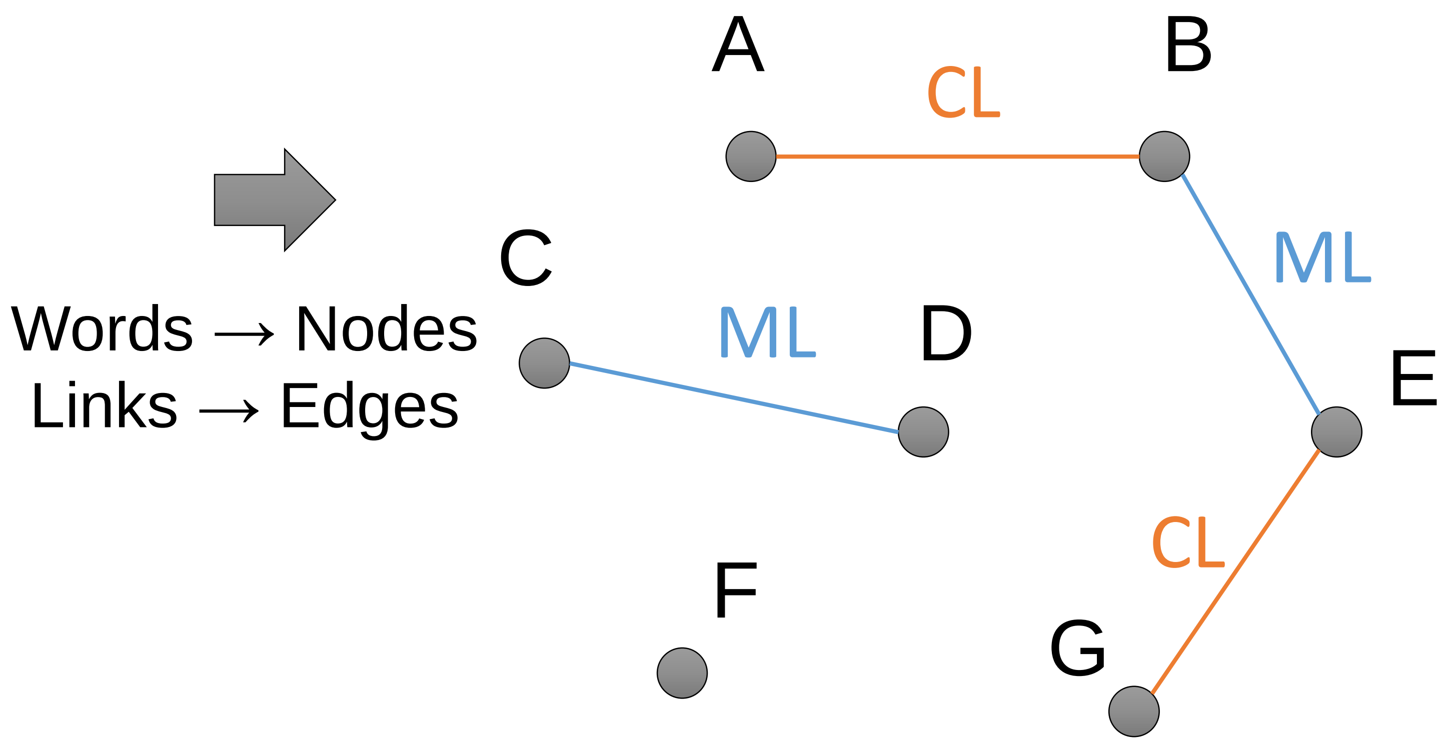
Algorithm to generate distributions in LDA-DF

1. Map links to a graph
2. Contract Must-Links
3. Extract the maximal independent sets (MIS)
4. Generate a distribution based on each MIS

Algorithm to generate distributions in LDA-DF

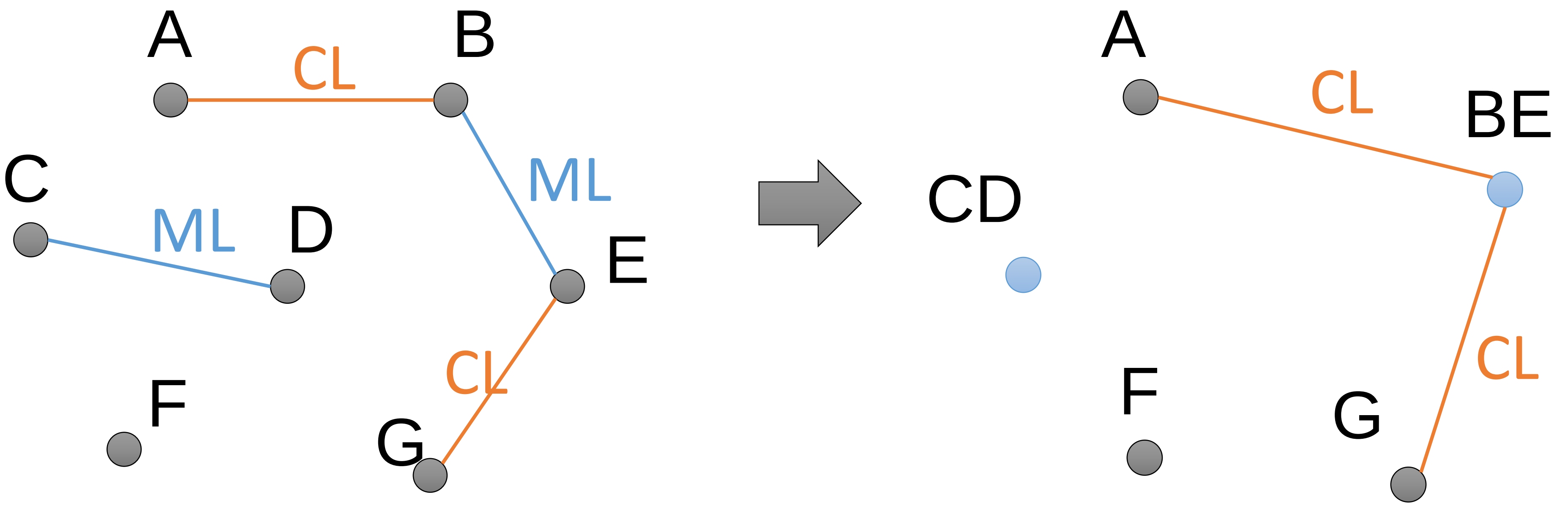
1. Map links to a graph
   - Any conjunction of links can be mapped to a graph

Cannot-Link(A,B) ∧ Cannot-Link(E,G) ∧ Must-Link(B,E) ∧ Must-Link(C,D)



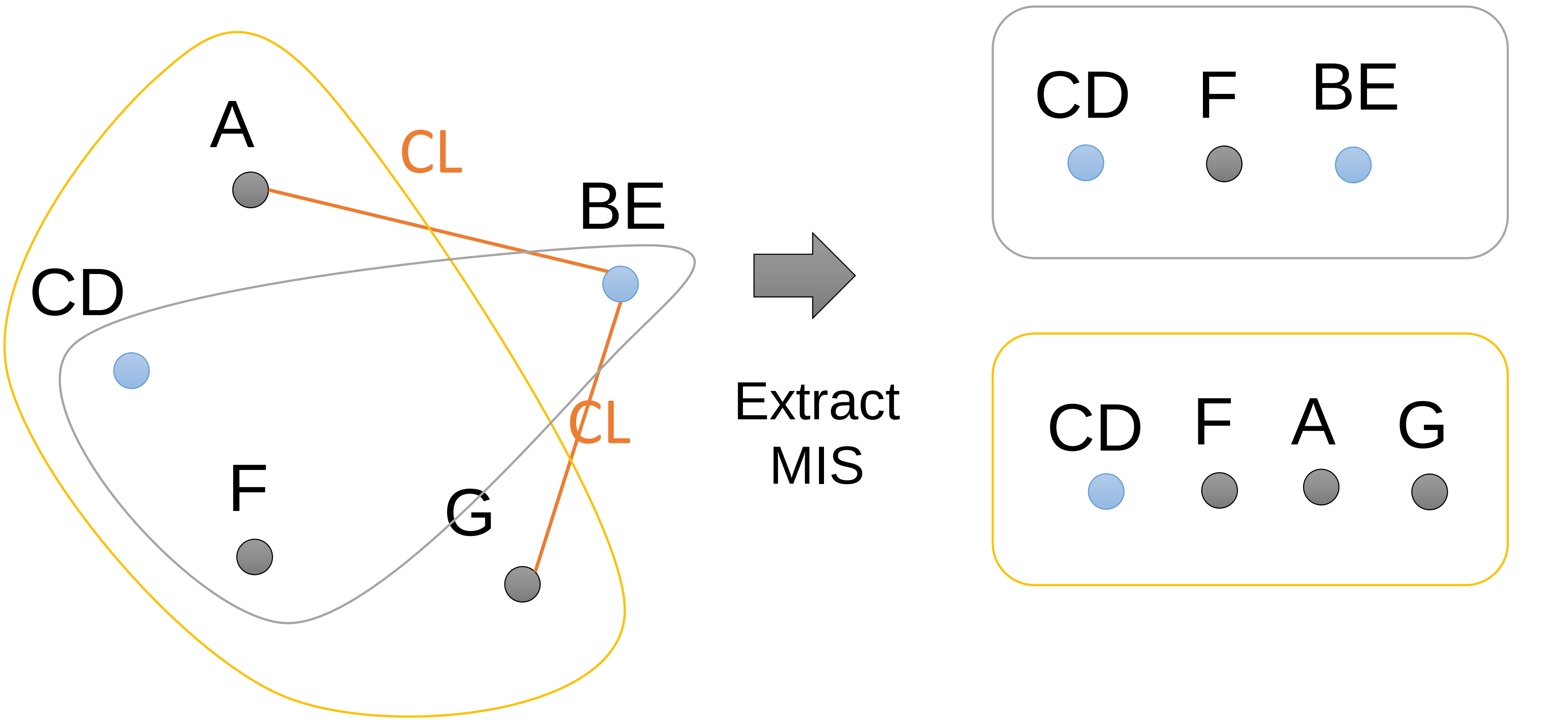
Algorithm to generate distributions in LDA-DF

2. Contract Must-Links
   - Regard two words on each Must-Link as one word



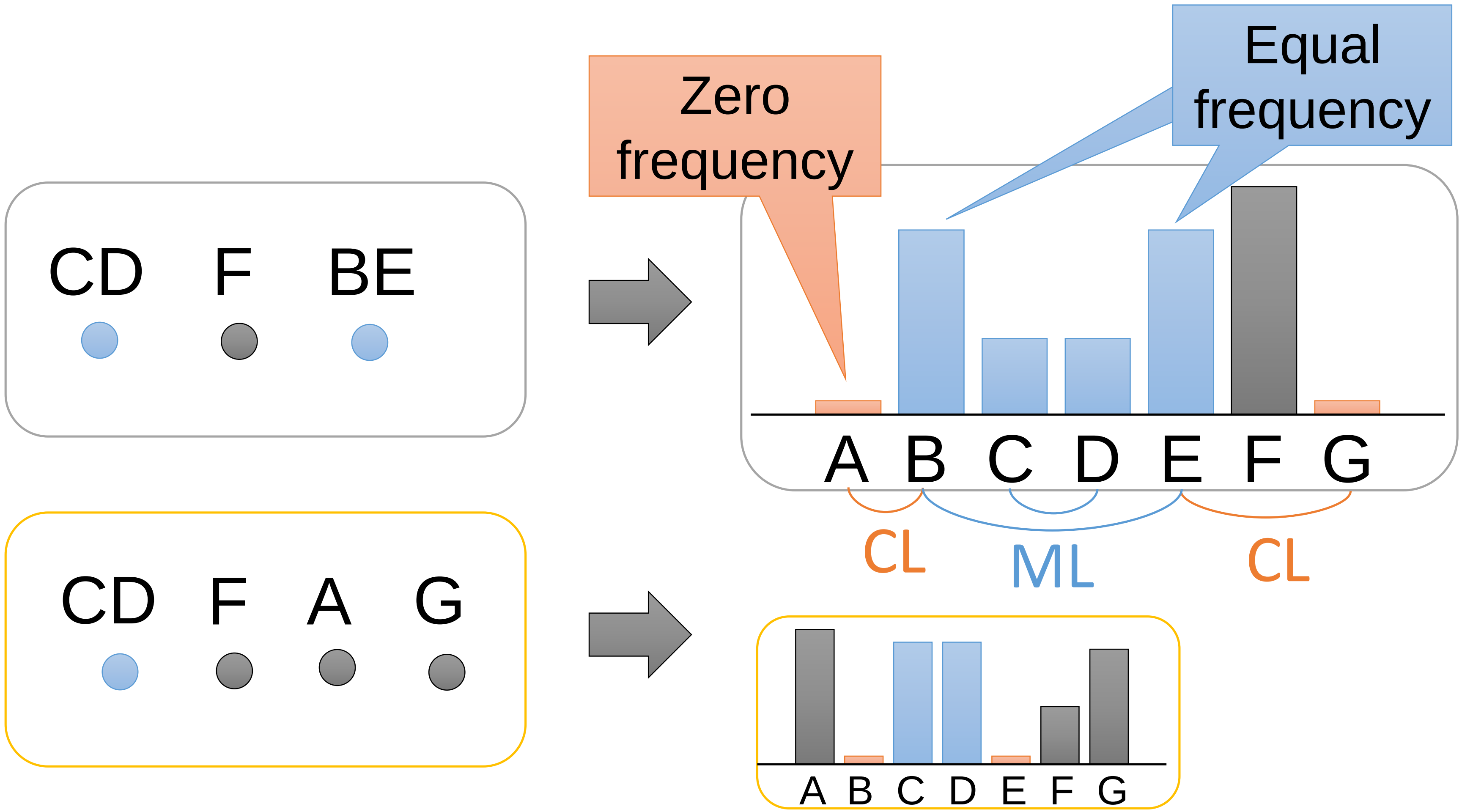
Algorithm to generate distributions in LDA-DF

3. Extract the maximal independent sets (MIS)
    - MIS = Maximal set of nodes without edges



Algorithm to generate distributions in LDA-DF

4. Generate a distribution based on each MIS
   - Equalize the frequencies of contracted words
   - Zero the frequencies of words **not** in the MIS



This work

- Algorithm to generate <u>logically</u> constrained distributions on LDA-DF
  - We can not apply the existing algorithm

$(\neg$ Cannot-Link(A,B)
$\vee$ Must-Link(A,C)$)$
$\wedge$ Cannot-Link(B,C)

Words → Nodes
Links → Edges

This constraint cannot be mapped to a graph

Negations

- Delete negations (¬) in a preprocessing stage
  - Weak negation: ¬Must-Link(A,B) = no constraint

    (A and B **need not** appear in the same topic)
  - Strong negation: ¬Must-Link(A,B) = Cannot-Link(A,B)

    (A and B **must not** appear in the same topic)

(¬Cannot-Link(A,B)
∨Must-Link(A,C))
∧Cannot-Link(B,C)

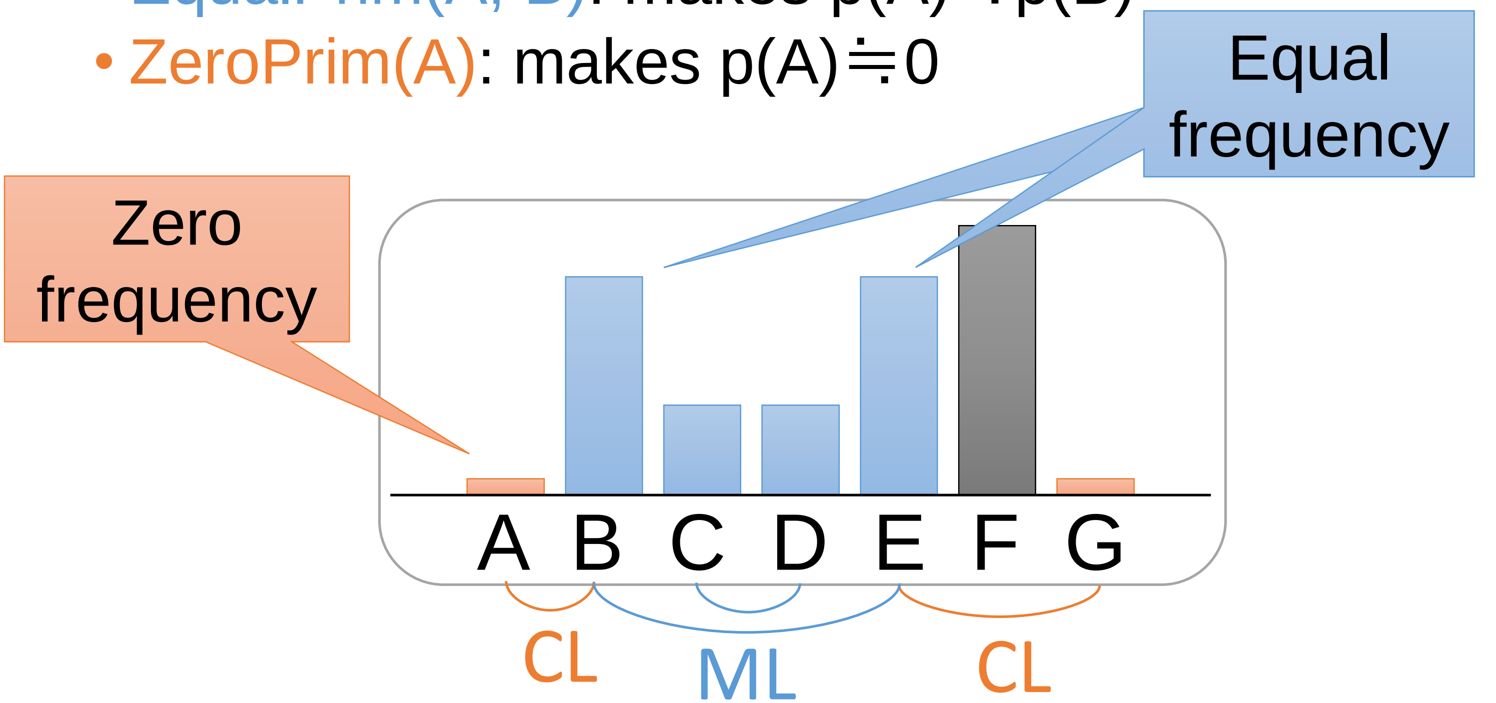
$\Rightarrow$

(Must-Link(A,B)
∨Must-Link(A,C))
∧Cannot-Link(B,C)

Focus only on conjunctions and disjunctions

Key observation for logical expressions

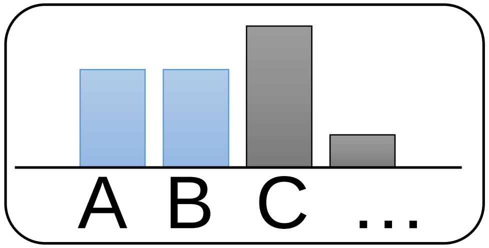
- Any constrained distribution is represented by a conjunctive expression by two primitives
  - EqualPrim(A, B): makes p(A)≒p(B)
  - ZeroPrim(A): makes p(A)≒0



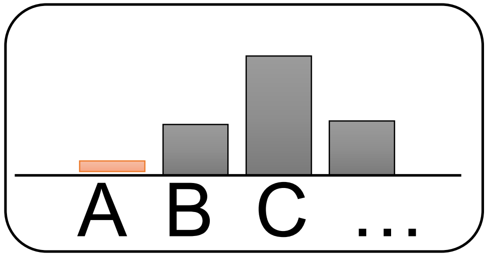
⇒ EqualPrim(B, E) ∧ EqualPrim(C, D) ∧ ZeroPrim(A) ∧ ZeroPrim(G)

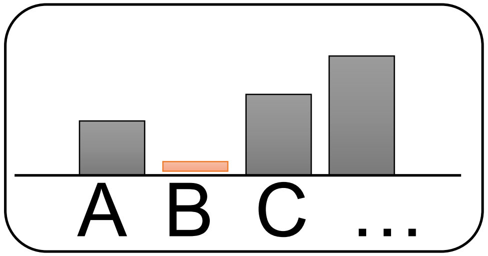
Substitution of links with primitives

- Must-Link(A,B) = EqualPrim(A,B)



- Cannot-Links(A,B) = ZeroPrim(A) $\vee$ ZeroPrim(B)





These two distributions satisfy Cannot-Link(A,B)

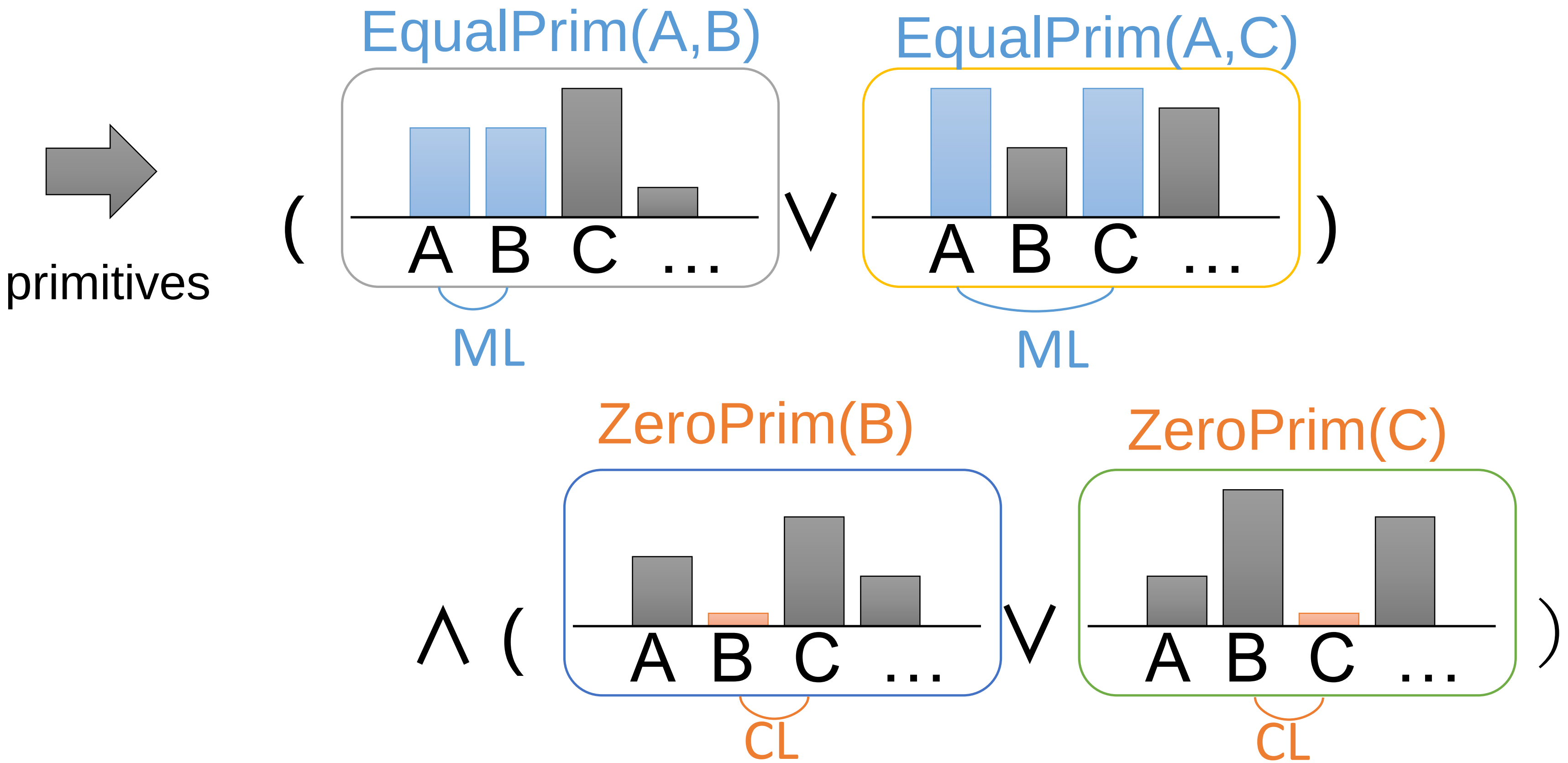
Proposed algorithm for logical expressions

1. Substitute links with primitives
2. Calculate the minimum disjunctive normal form (DNF) of the primitives
3. Generate distributions for each conjunction of the DNF

Proposed algorithm for logical expressions

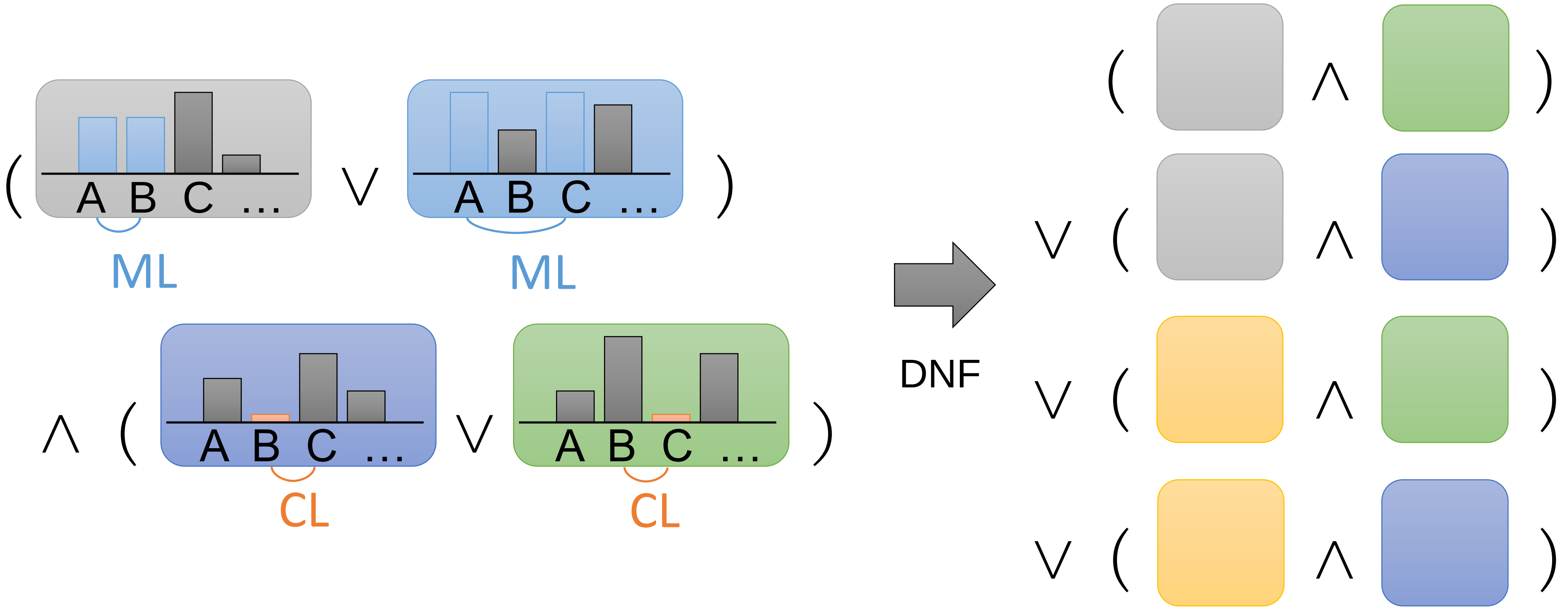
1. Substitute links with primitives

(Must-Link(A,B) ∨ Must-Link(A,C))
∧ Cannot-Link(B,C)



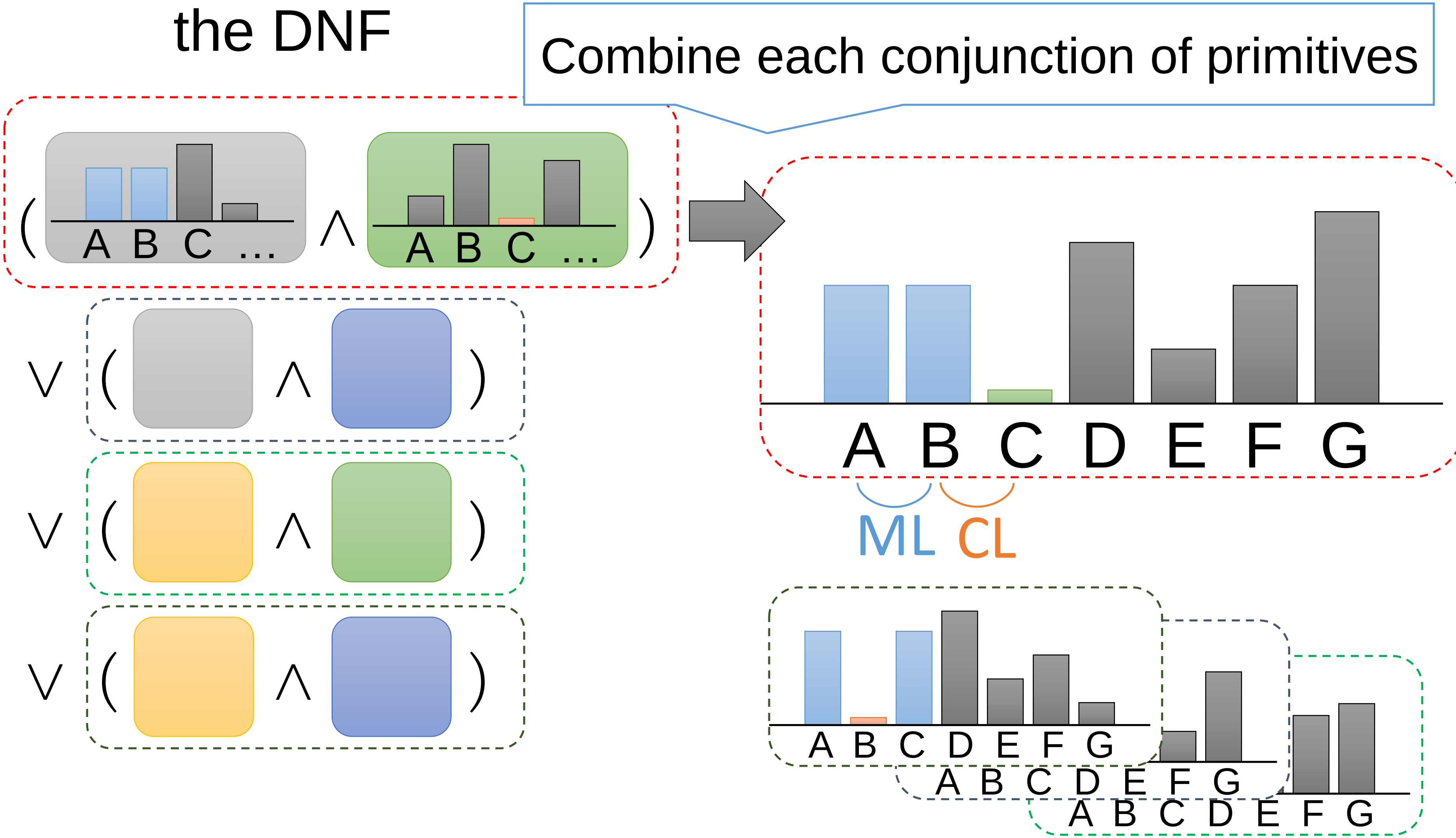
Proposed algorithm for logical expressions

2. Calculate the minimum disjunctive normal form (DNF) of the primitives
   - DNF = Disjunction of conjunctions of primitives



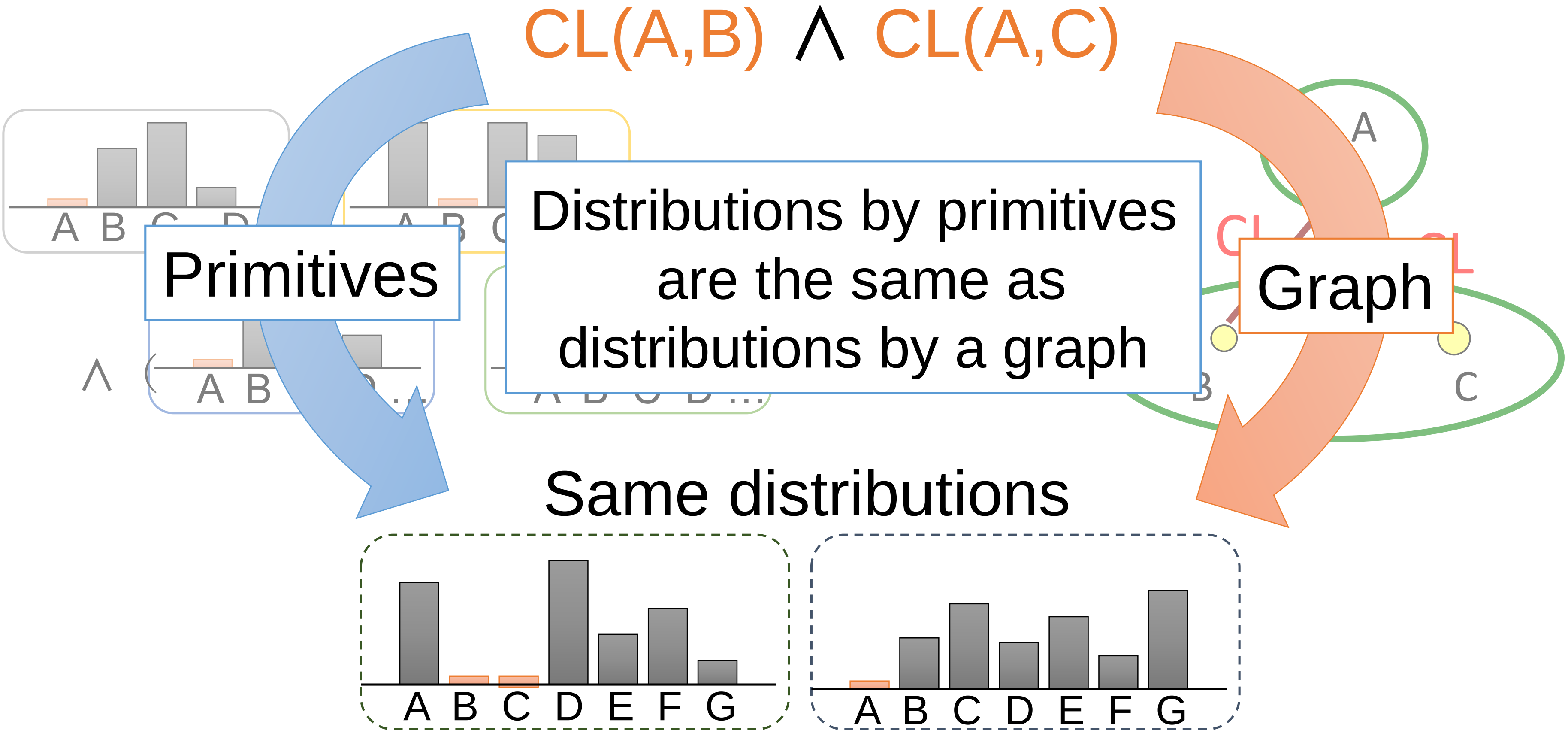
Proposed algorithm for logical expressions

3. Generate distributions for each conjunction of the DNF



Correctness of our method

- [Theorem] Our method and the existing method are asymptotically equivalent w.r.t. <u>conjunctive expressions of links</u>



Customization of new links

- Isolate-Link (ISL)
  - $X_1,\ldots,X_n$ do not appear (nearly) (Remove unnecessary words and stop words）

$$\mathrm{ISL}(X_1,\ldots,X_n) \;=\; \bigwedge_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{ZeroPrim}(X_i)$$

- Imply-Link (IL)
  - B appears if A appears in a topic (A→B) (Use when B has multiple meanings)

$$\mathrm{IL}(A,B) \;=\; \mathrm{EqualPrim}(A,B) \vee \mathrm{ZeroPrim}(A)$$

- Extended Imply-Link (XIL)
  - Y appears if $X_1,\ldots,X_n$ appear in a topic ($X_1,\ldots,X_n \rightarrow Y$)

$$\mathrm{XIL}(X_1,\ldots,X_n,Y) \;=\; \bigwedge_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{EqualPrim}(X_i,Y) \vee \bigvee_{i=1}^{n} \mathrm{ZeroPrim}(X_i)$$

Interactive topic analysis

- Movie review corpus (1000 reviews) [Pang&Lee ACL2004]
  - No constraints

Topic	High frequency words
?	**have** give night **film** turn performance year mother take out
?	**not** life **have** own first only family tell yet moment even
?	**movie have n’t** get good **not** see know just other time make
?	**have** black scene tom death die joe ryan man final private
?	**film have n’t not** make out well see just very watch even
?	**have film** original new never more evil **n’t** time power
…	…

All topics are unclear

Interactive topic analysis

- Movie review corpus (1000 reviews)
  - Isolate-Link(have, film, movie, not, n’t)
    - Remove specified words as well as related unnecessary words

Topic	High frequency words
(Isolated)	**have film movie not** good make **n’t** character see more get
?	**star war trek** planet effect special lucas jedi science
Comedy	comedy funny laugh school hilarious evil power bulworth
Disney	disney voice mulan animated song feature tarzan animation
Family	life love family mother woman father child relationship
Thriller	truman murder killer death thriller carrey final detective
…	…

“Star Wars” and “Star Trek” are merged, although most topics are clear

Interactive topic analysis

- Movie review corpus (1000 reviews)
  - Isolate-Link(have, film, movie, not, n't)
    $\wedge$ Cannot-Link(jedi, trek)

Dared to select "jedi" since "star" and "war" are too common

Topic	High frequency words
(Isolated)	**have film movie not** make good **n't** character see more get
Star Wars	**star war** lucas effect **jedi** special matrix menace computer
Comedy	funny comedy laugh get hilarious high joke humor bob smith
Disney	disney truman voice toy show animation animated tarzan
Family	family father mother boy child son parent wife performance
Thriller	killer murder case lawyer man david prison performance
…	…

"Star Trek" disappears, altough "Star Wars" is obtained

Interactive topic analysis

- Movie review corpus (1000 reviews)
  - Isolate-Link(have, film, movie, not, n't)
    $\wedge$ Cannot-Link(jedi, trek)
    $\wedge$ (Must-Link(star, jedi) $\vee$ Must-Link(star, trek))

Topic	High frequency words
(Isolated)	**have film movie not** make good **n't** character see more get
Star Wars	**star war** toy **jedi** menace phantom lucas burton planet
Star Trek	alien effect **star** science special **trek** action computer
Comedy	comedy funny laugh hilarious joke get ben john humor fun
Disney	disney voice animated mulan animation family tarzan shrek
Family	life love family man story child woman young mother
Thriller	scream horror flynt murder killer lawyer death sequel case
…	…

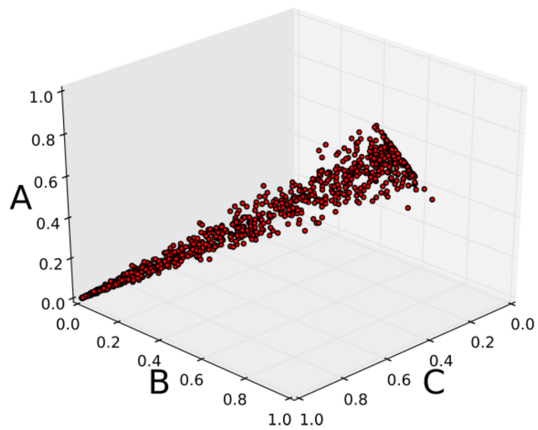
We obtained "Star Wars" and "Star Trek" appropriately

Conclusion

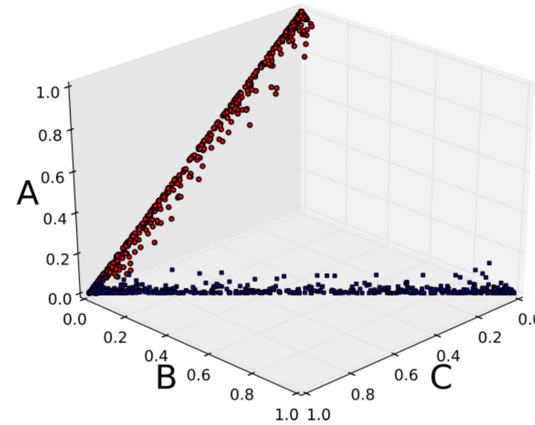
- Simple algorithm for logical constraints on words for topic modeling
  - Must-Link(A,B)： A and B appear in the same topic
  - Cannot-Link(A,B)： A and B do not appear in the same topic
- Theorem for the correctness of the algorithm
- Customization of new links
  - Isolate-Link(X_1, …, X_n): X_1, …, X_n disappear
  - Imply-Link(A, B): B appears if A appears in a topic
- Future Work
  - Comparative experiments on real corpora

Thank you for your attention

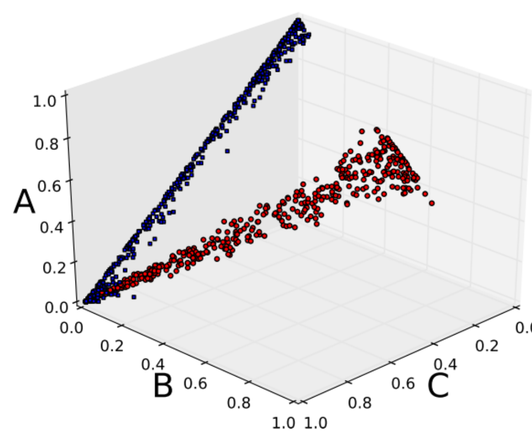
Appendix: Visualization of Priors



(a) $ML(A, B)$



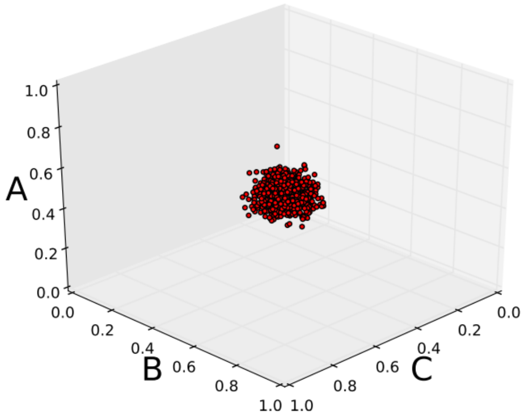
(b) $CL(A, B)$



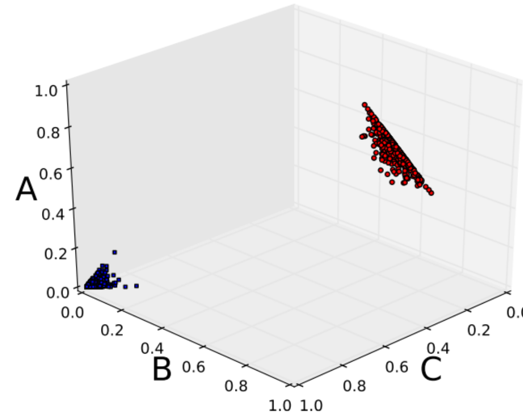
(c) $IL(B, A)$

ML = Must-Link, CL = Cannot-Link, IL = Imply-Link

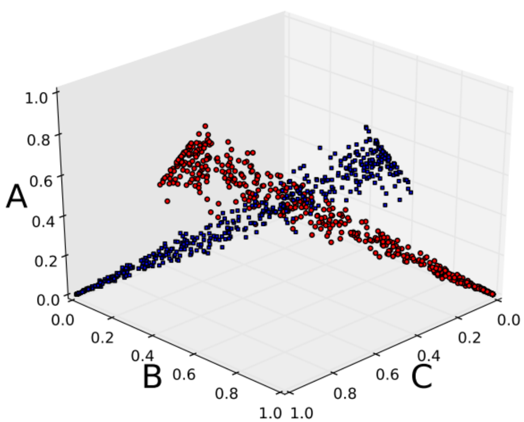
Appendix: Visualization of Priors



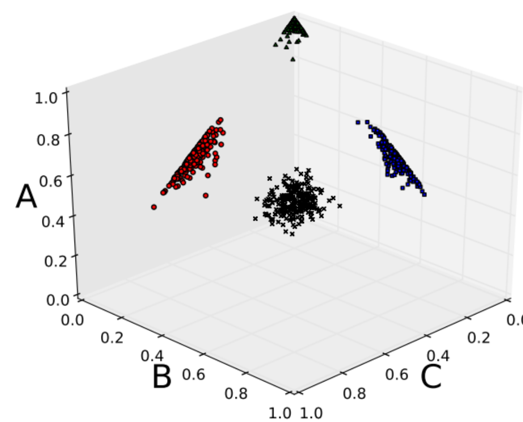
(a) $ML(A, B) \wedge ML(A, C)$



(b) $ML(A, B) \wedge CL(B, C)$



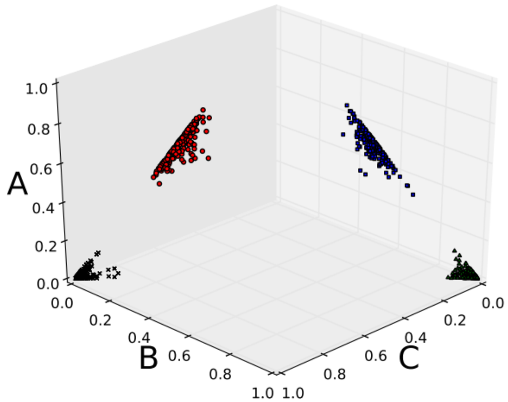
(c) $ML(A, B) \vee ML(A, C)$



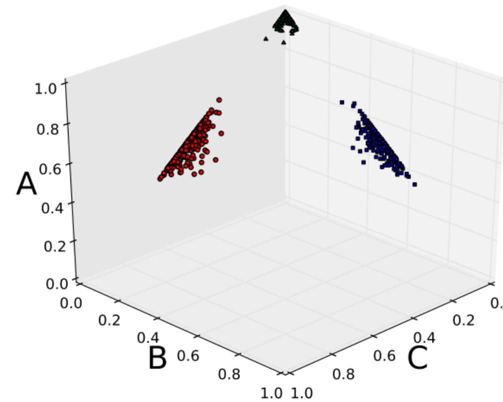
(d) $IL(B, A) \wedge IL(C, A)$

ML = Must-Link, CL = Cannot-Link, IL = Imply-Link

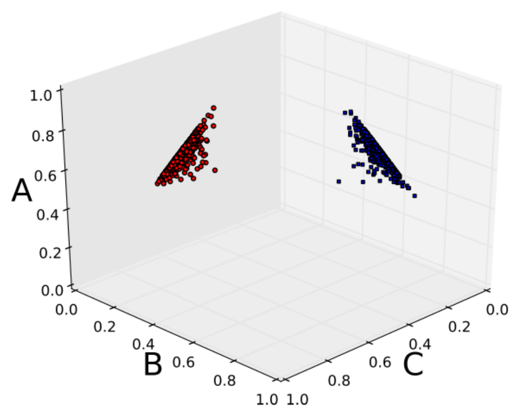
Appendix: Visualization of Priors



(e) $(ML(A,B) \vee ML(A,C)) \wedge CL(B,C)$



(f) $IL(B,A) \wedge IL(C,A) \wedge CL(B,C)$



(g) $ML(C,A) \veebar ML(C,B)$

ML = Must-Link, CL = Cannot-Link, IL = Imply-Link