max planck institut informatik

Max Planck Institute for Software Systems

People on Drugs : Credibility of User Statements in Health Forums

Subhabrata Mukherjee [1] **Gerhard Weikum** [1]
Cristian Danescu-Niculescu-Mizil [2]

[1] **Max Planck Institute for Informatics**

[2] **Max Planck Institute for Software Systems**

KDD 2014

August 25, 2014

Motivation: Internet as a healthcare resource

59% of US population use internet for health information

[Pew Research Center Report, 2013]

Half of US physicians rely on online resources

[IMS Health Report, 2014]

This work:

Credibility of user-generated online health information

Motivation: Internet as a healthcare resource

59% of US population use internet for health information

[Pew Research Center Report, 2013]

Half of US physicians rely on online resources

[IMS Health Report, 2014]

This work:

Credibility of user-generated online health information

Posts from Healthboards.com

"My girlfriend always gets a bad dry skin, rash on her upper arm, cheeks, and shoulders when she is on [Depo]. . . ."

"I have had no side effects from [Depo] (except ...), but otherwise no rashes. She should see her gyno. She may be allergic to something"

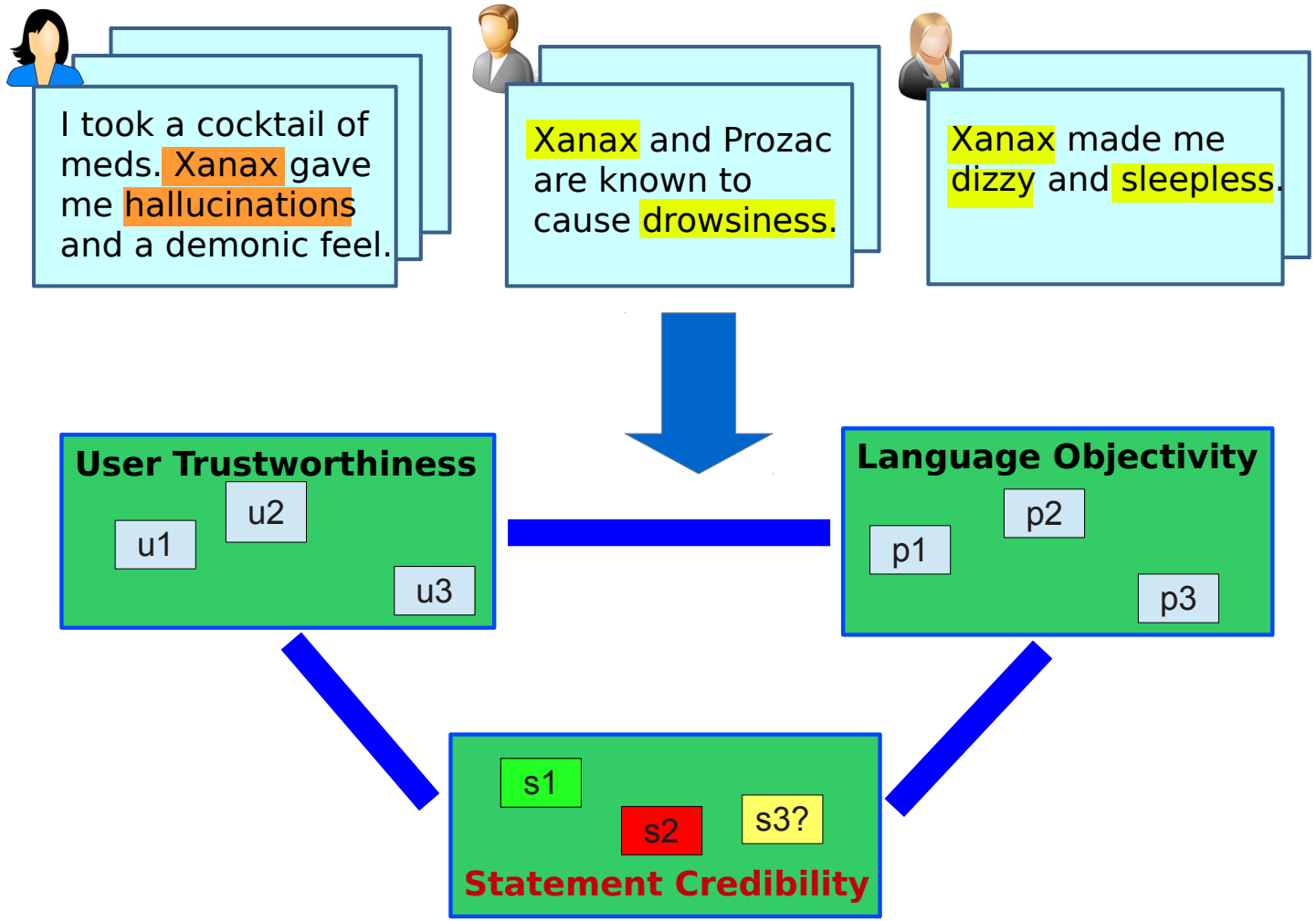
Posts from Healthboards.com

"My girlfriend always gets a bad dry skin, rash on her upper arm, cheeks, and shoulders when she is on [Depo]. . . ."

"I have had no side effects from [Depo] (except ...), but otherwise no rashes. She should see her gyno. She may be allergic to something"

Our Intuition

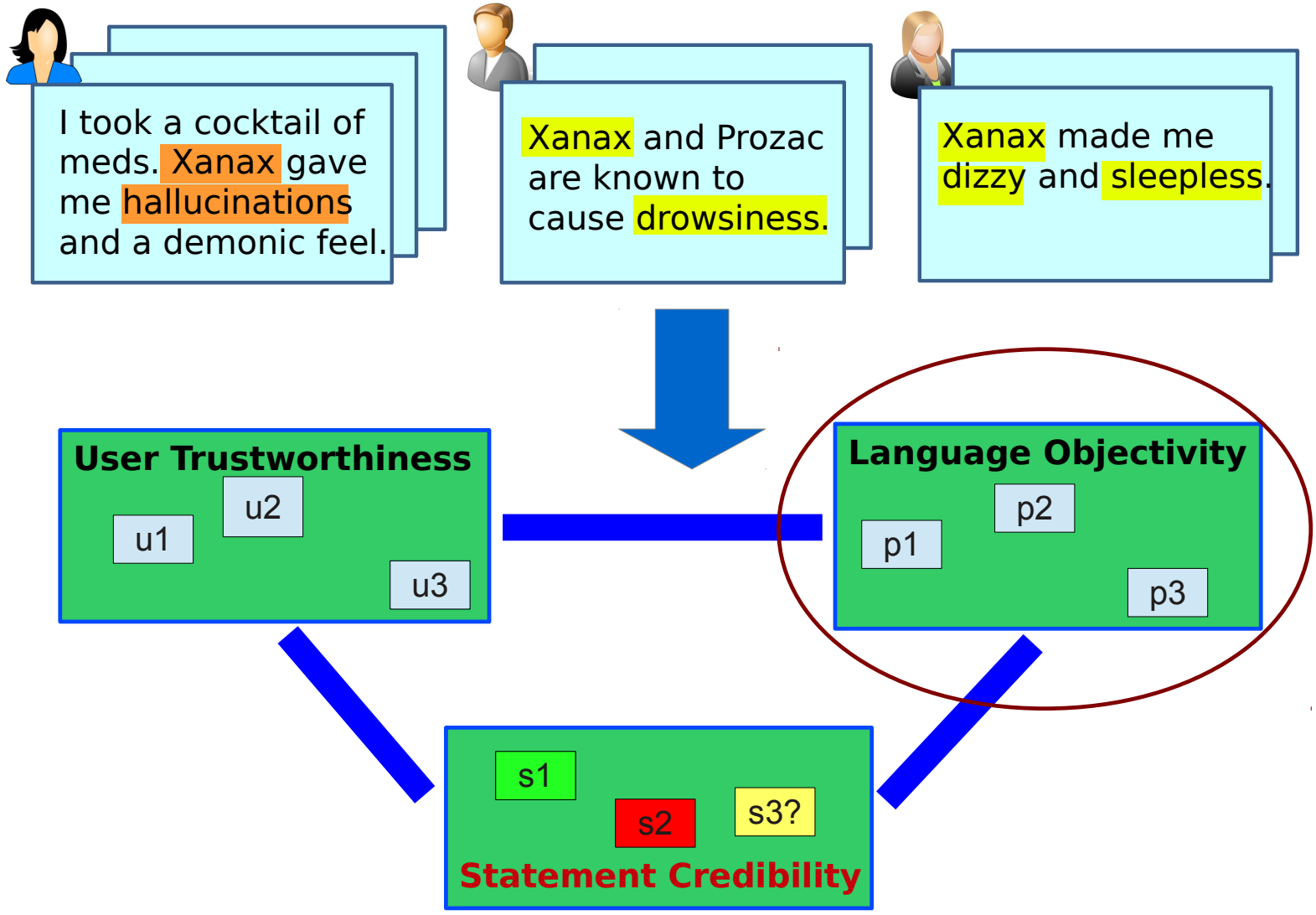
Users, *language* and *credibility* influence each other



Trustworthy users write *credible* posts

Agree with each other on *credible* statements

Our Intuition



Language: Stylistic Features

"I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience ?"

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, *etc.*

Language: Stylistic Features

> “I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience ?”

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, *etc.*

Language: Stylistic Features

> “I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience ?”

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, *etc.*

Language: Stylistic Features

> “I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience ?”

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, *etc.*

Language: Stylistic Features

"I heard Xanax can have pretty bad side-effects. You may have peeling of skin, and apparently some friend of mine told me you can develop ulcers in the lips also. If you take this medicine for a long time then you would probably develop a lot of other physical problems. Which of these did you experience ?"

Usage of modals, indefinite determiner, conditional, probabilistic adverb, question particle, *etc.*

Language: Stylistic Features

"Depo is very dangerous as a birth control and has too many long term side-effects like reducing bone density. Hence, I will never recommend anyone using this as a birth control. Some women tolerate it well but those are the minority. Most women have horrible long lasting side-effects from it."

Uses inferential conjunction, modal, definite determiners, *etc.*

Language: Stylistic Features

"Depo is very dangerous as a birth control and has too many long term side-effects like reducing bone density. Hence, I will never recommend anyone using this as a birth control. Some women tolerate it well but those are the minority. Most women have horrible long lasting side-effects from it."

Uses inferential conjunction, modal, definite determiners, *etc.*

Language: Stylistic Features

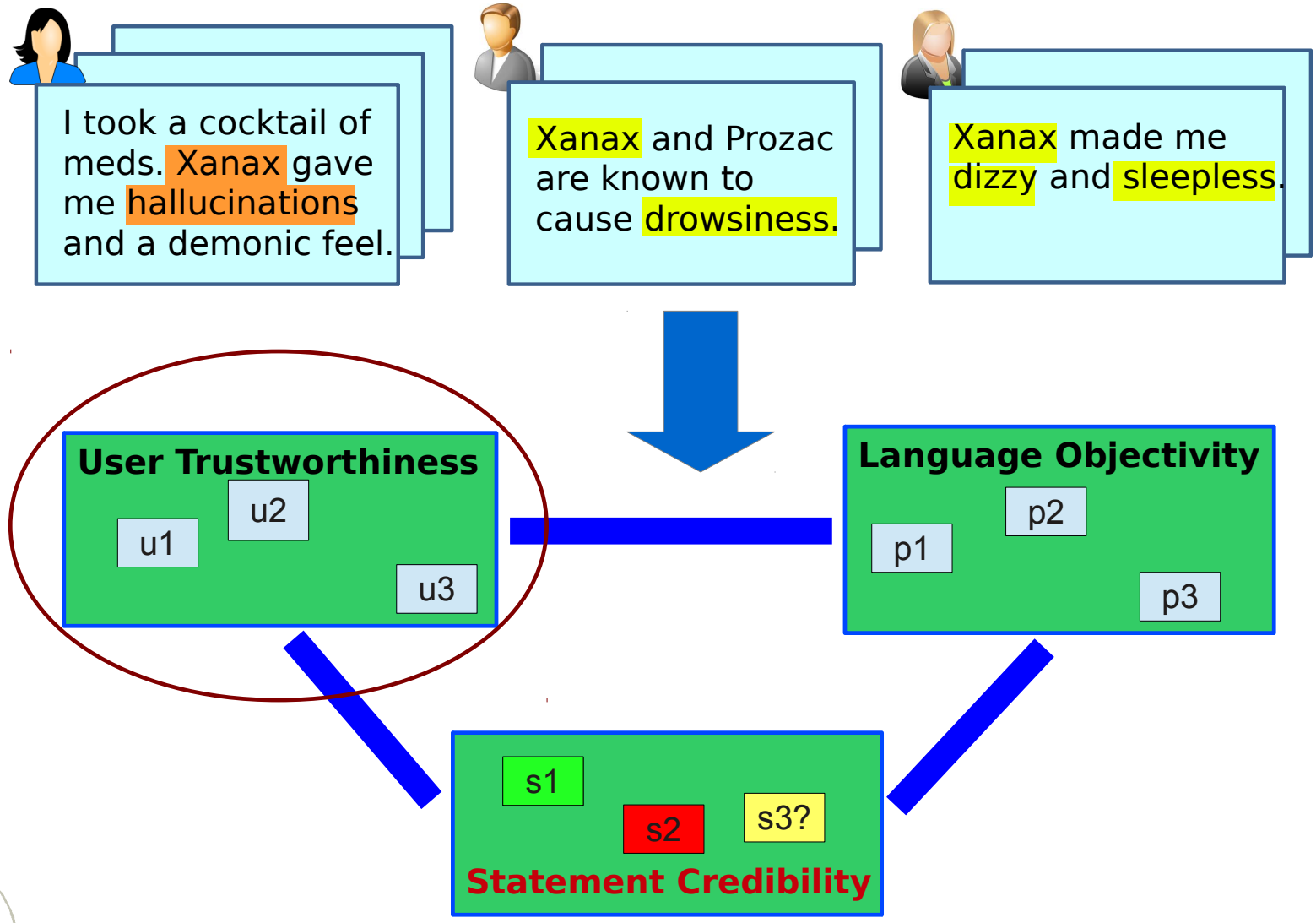
"Depo is very dangerous as a birth control and has too many long term side-effects like reducing bone density. Hence, I will never recommend anyone using this as a birth control. Some women tolerate it well but those are the minority. Most women have horrible long lasting side-effects from it."

Uses inferential conjunction, modal, definite determiners, *etc.*

Language: Objectivity

"I started Cymbalta, but now I'm having a panic attack or an allergic reaction. I have a hardcore burning sensation in my chest and warm sensations all over. It's like my body can't decide whether it wants to be cold or hot. I feel if I close my eyes I'll lose control, go crazy and pass out."

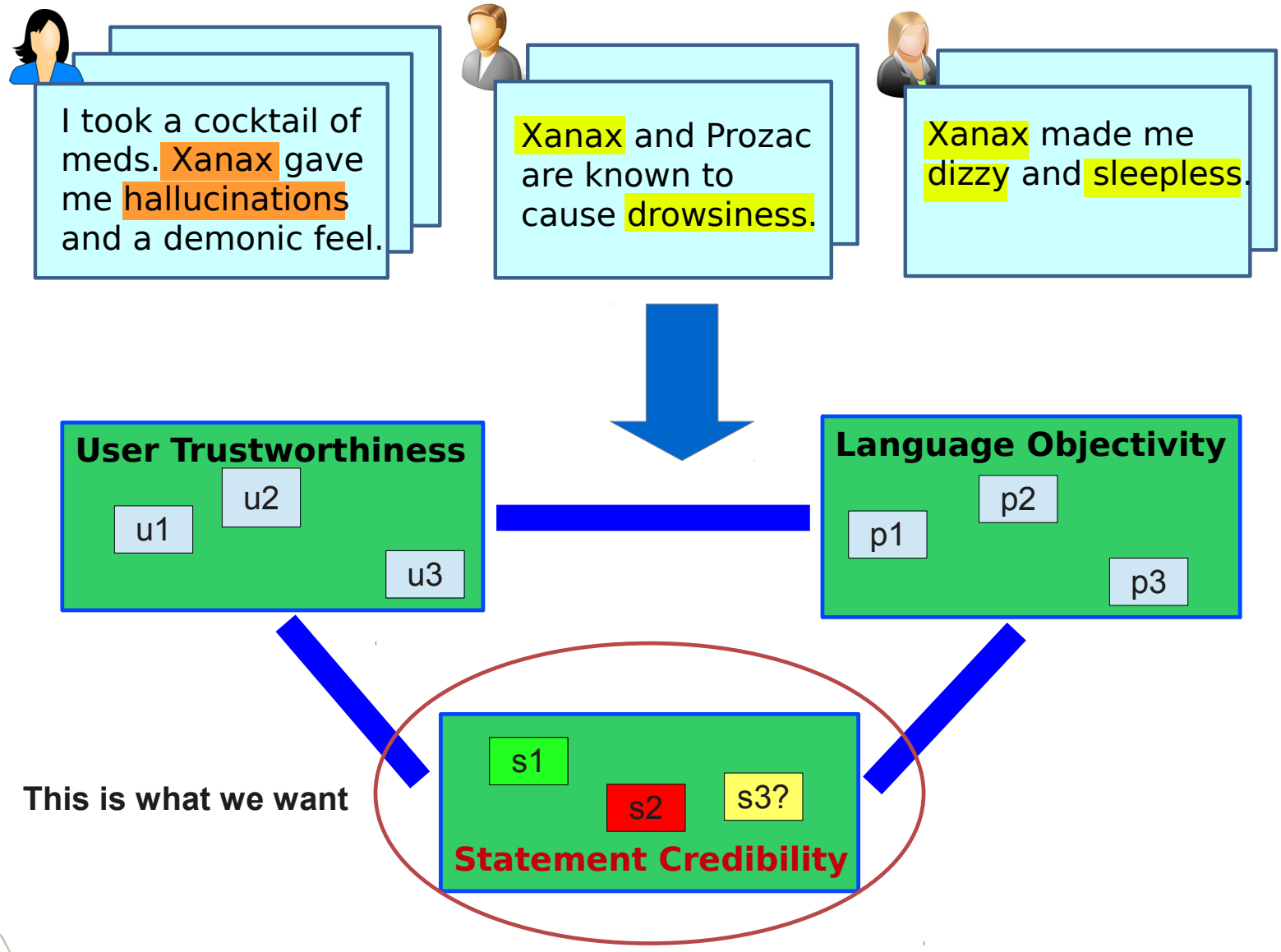
Our Intuition



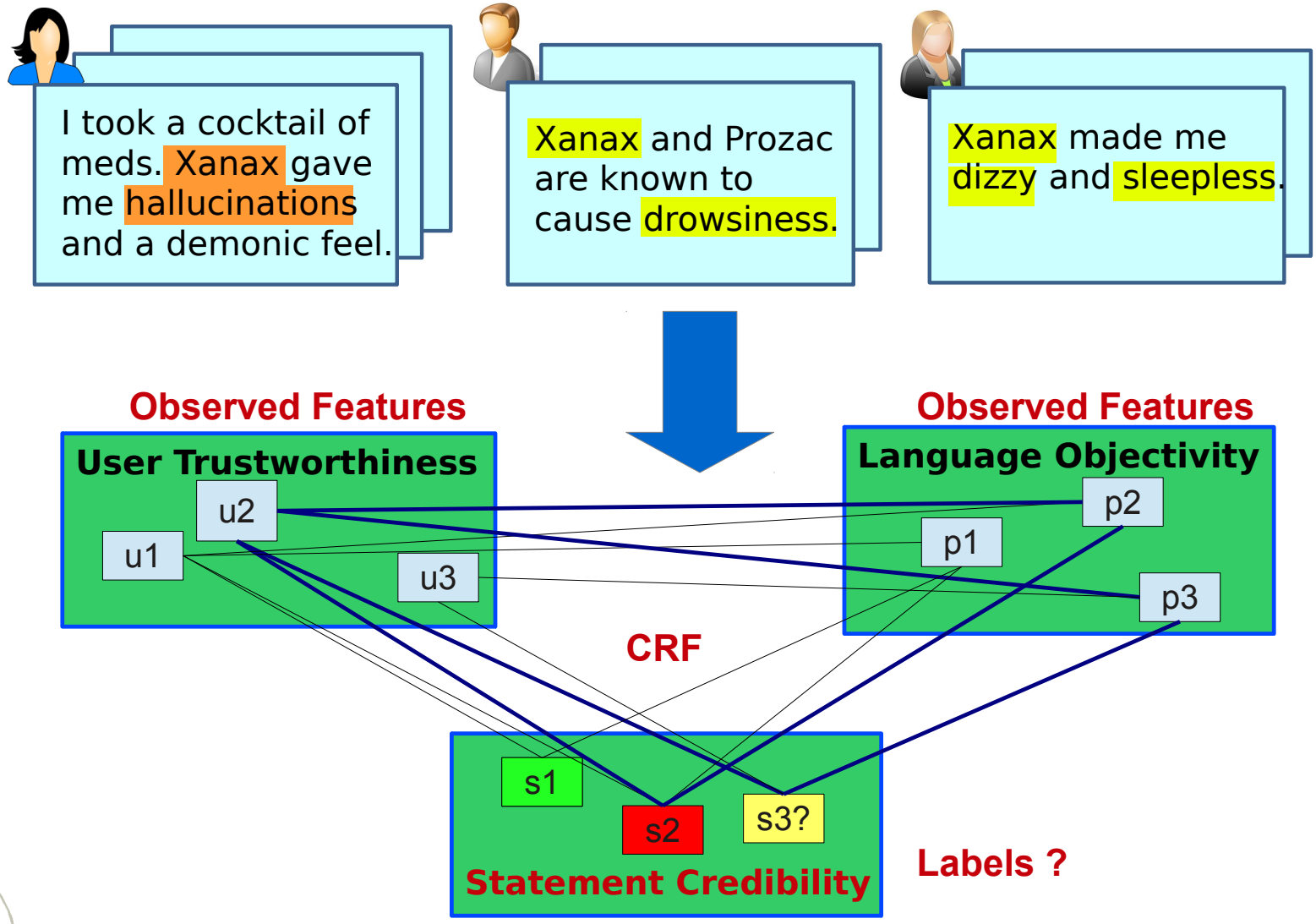
User Features

- User demographic features like *age, gender, location*
- Engagegement features like *number of posts, questions, answers, thanks*
- User post properties like *avg. post length*

Objective


I took a cocktail of meds. Xanax gave me hallucinations and a demonic feel.
Xanax and Prozac are known to cause drowsiness.
Xanax made me dizzy and sleepless.
User Trustworthiness
u1
u2
u3
Language Objectivity
p1
p2
p3
s1
s2
s3?
Statement Credibility
This is what we want

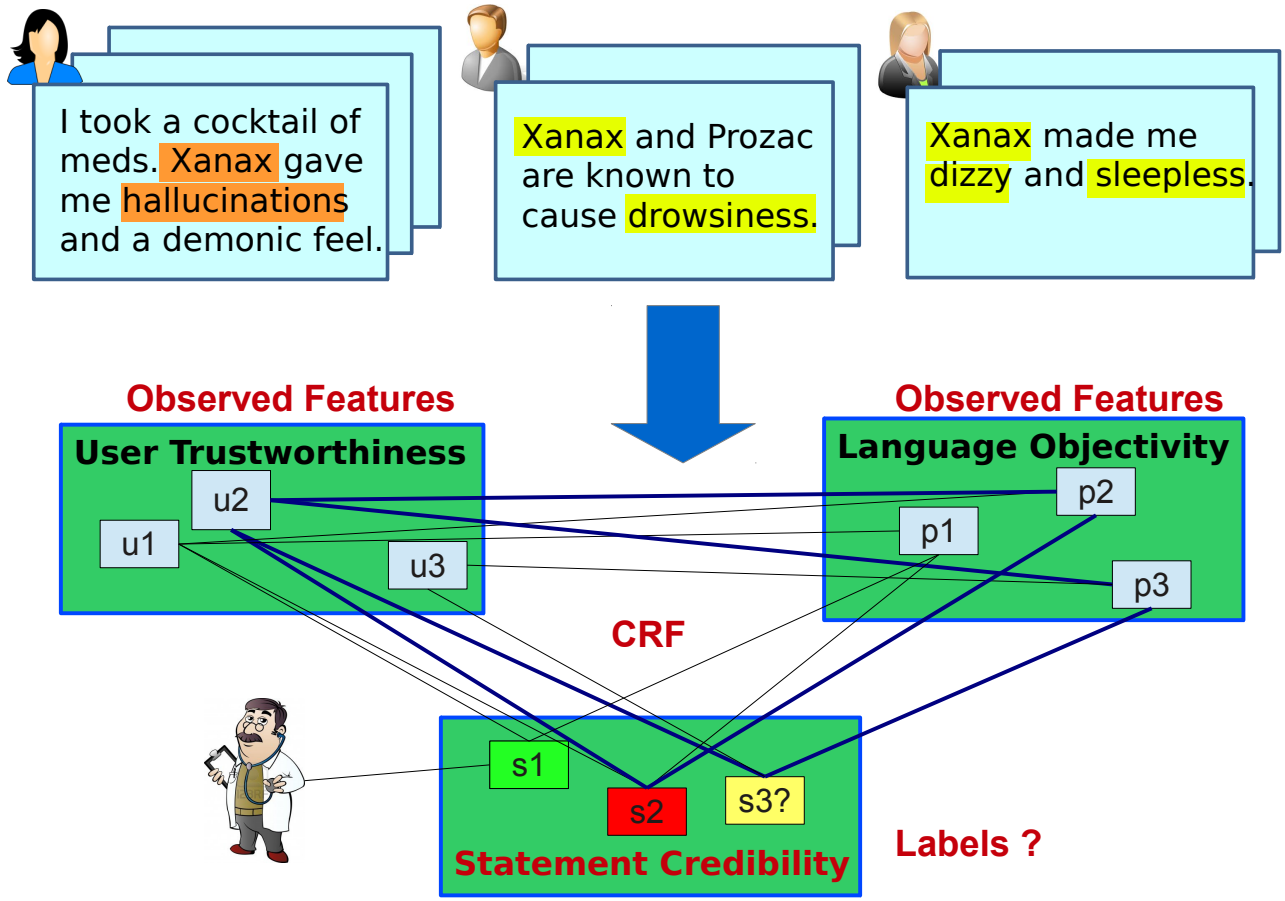
Probabilistic Inference: CRF



Predict the *most likely label assignment* of statements

Semi Supervised Learning

Protects against users conveying misinformation using confident and objective language



Expert stated side-effects of drugs from MayoClinic portal

Semi-Supervised CRF (Sketch)

User Trustworthiness

u_1 u_2 u_3

Language Objectivity

p_1 p_2 p_3

s_1 s_2?

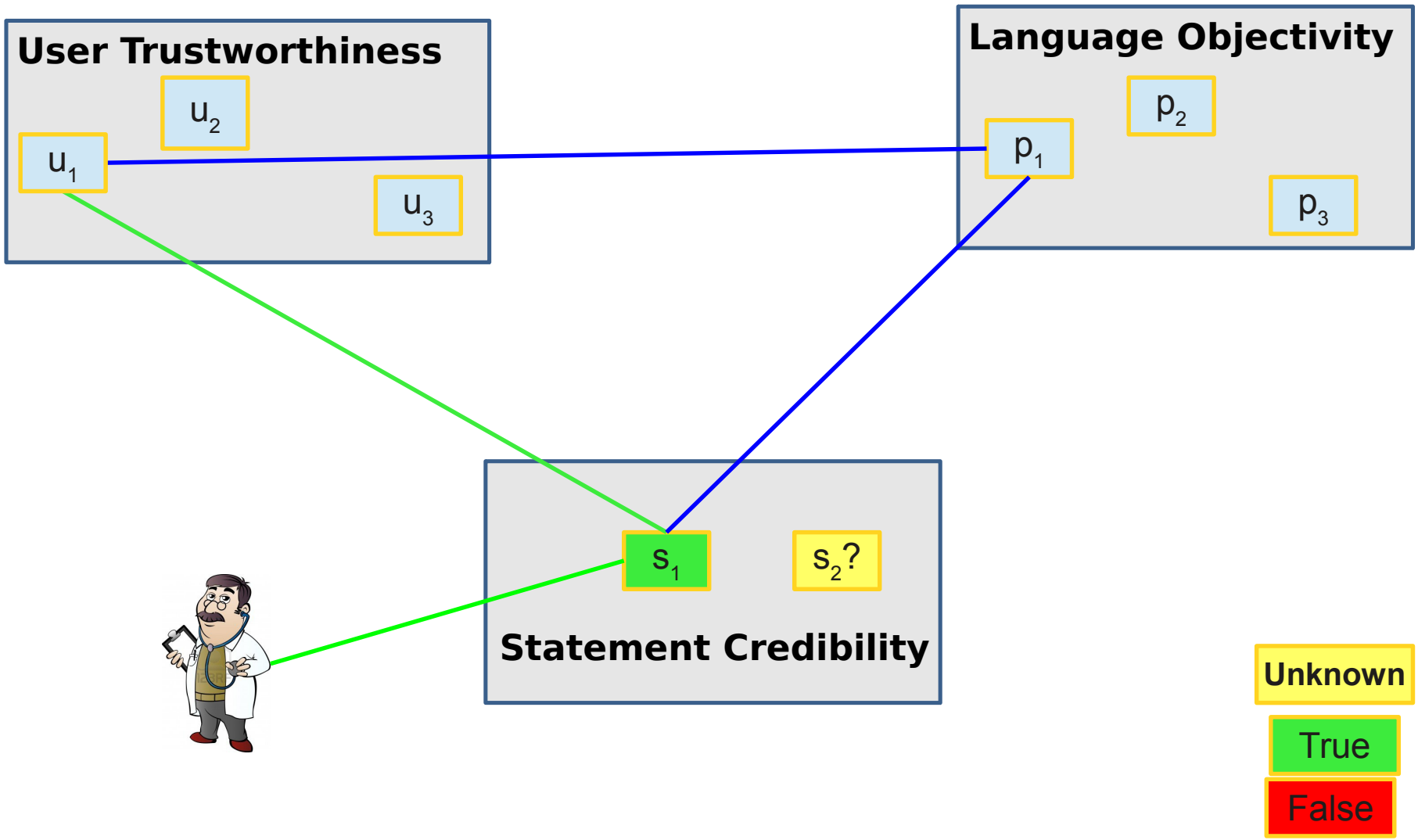
Statement Credibility

Unknown

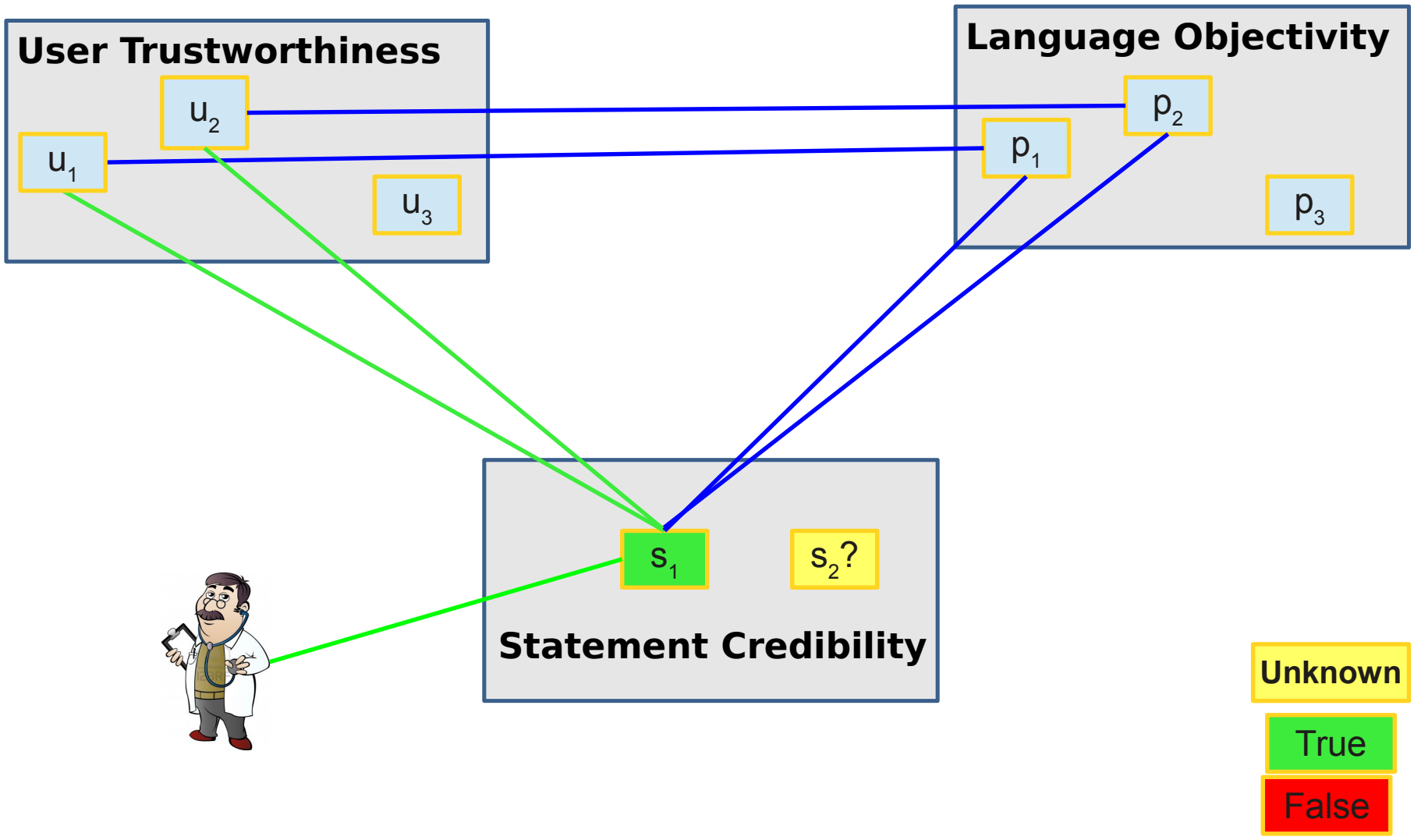
True

False

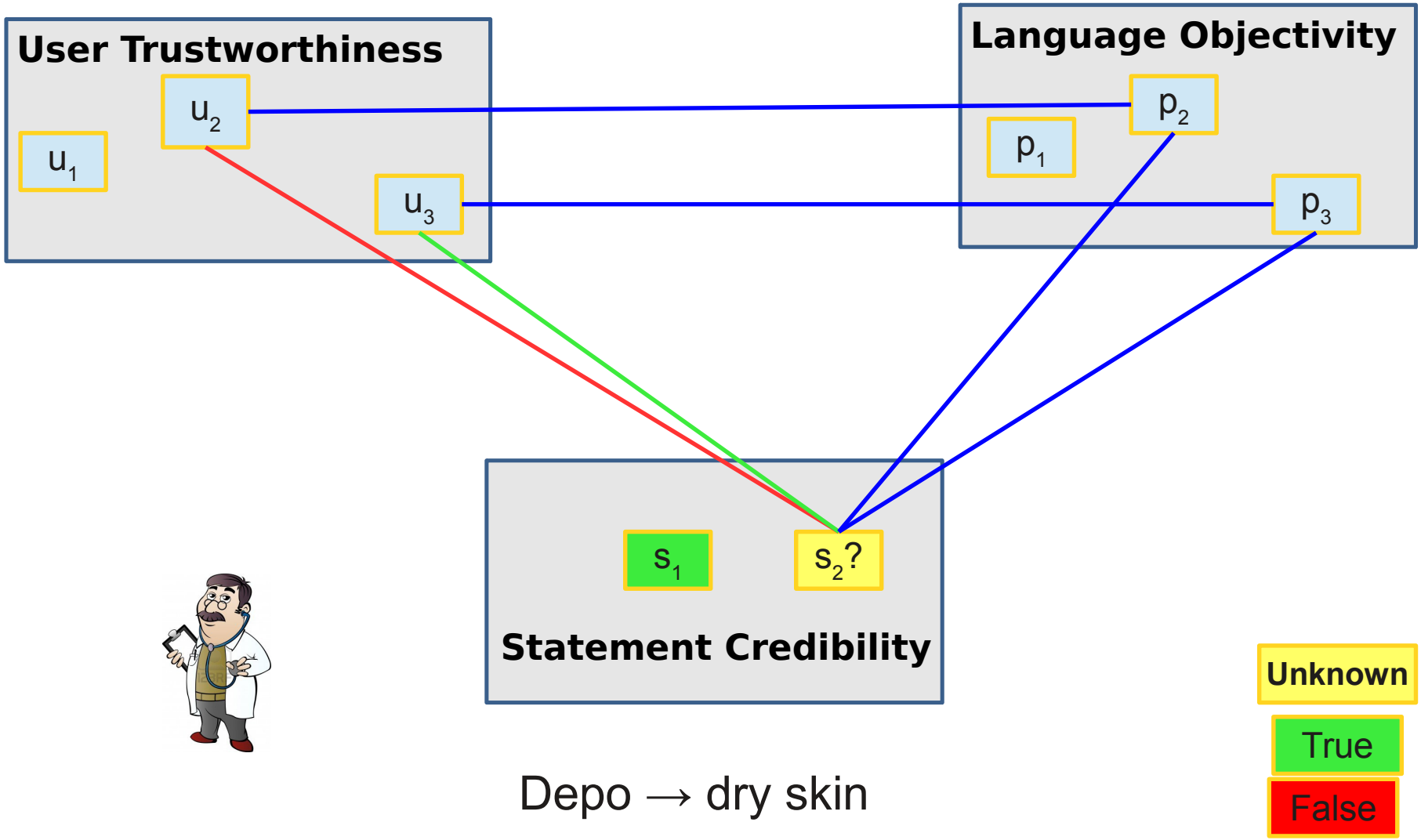
Semi-Supervised CRF (Sketch)



Semi-Supervised CRF (Sketch)

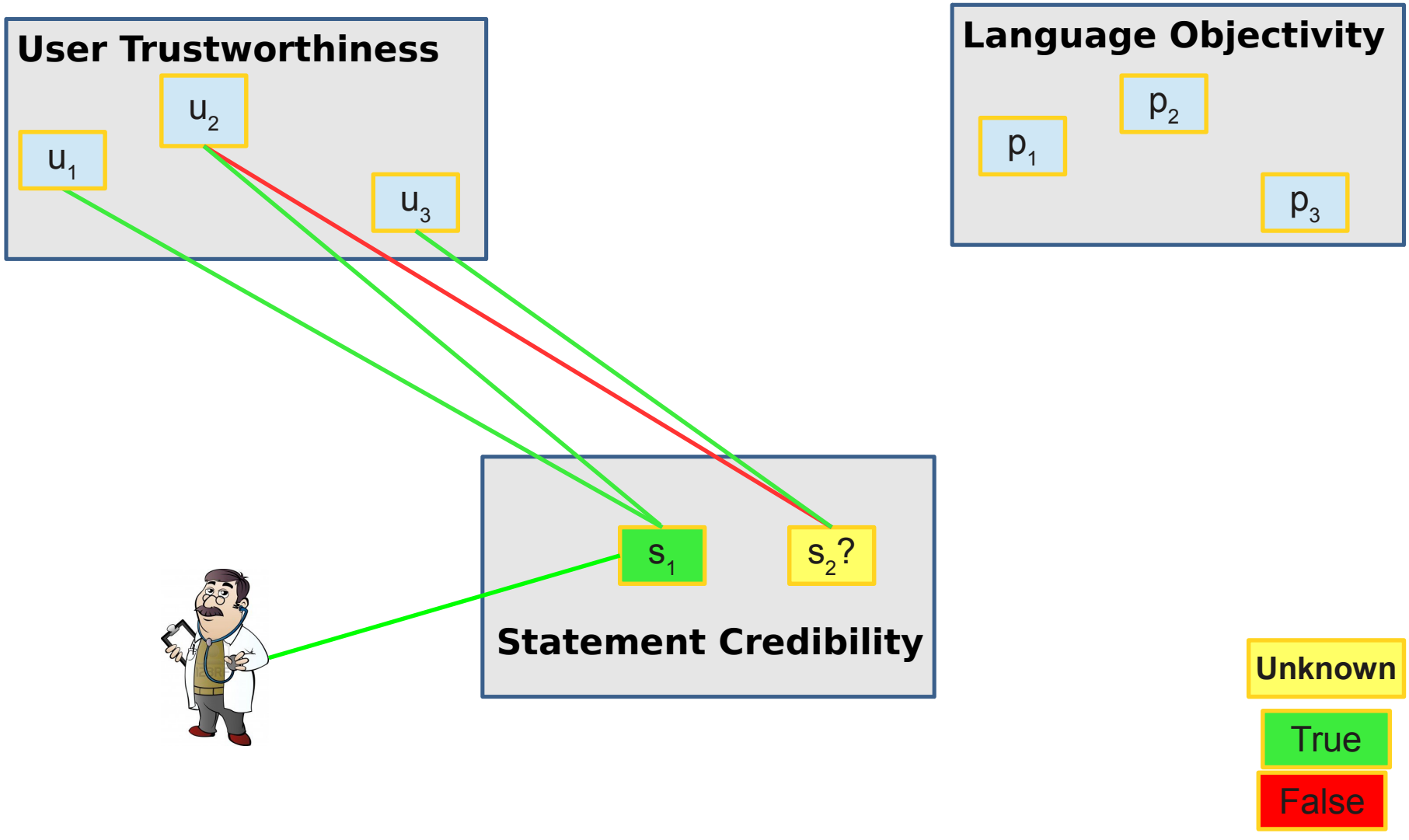


Semi-Supervised CRF (Sketch)

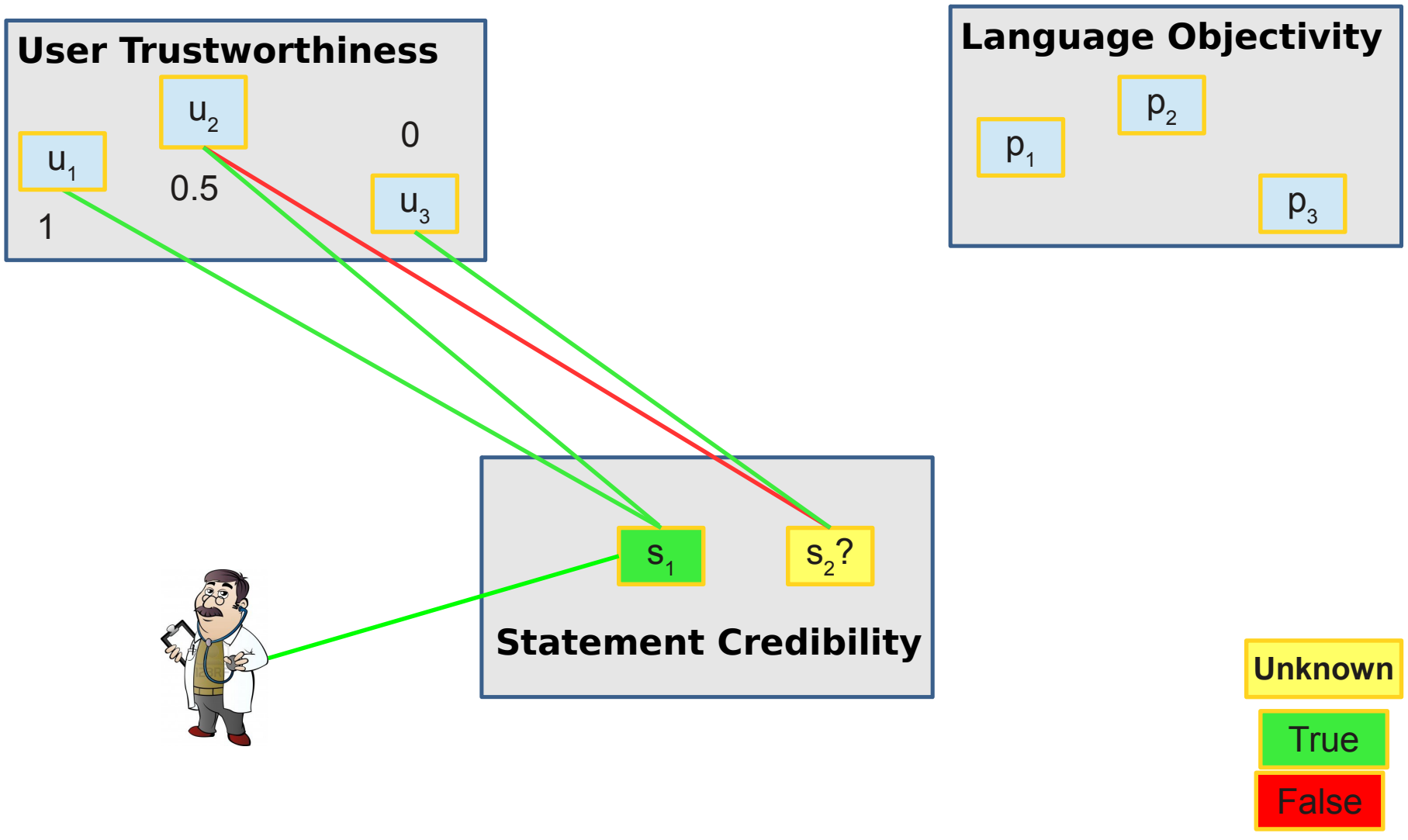


Depo → dry skin

1. Estimate user trustworthiness : $t_k = \frac{\sum_i \mathbb{1}_{S_{i,k}=\text{True}}}{|S_k|}$

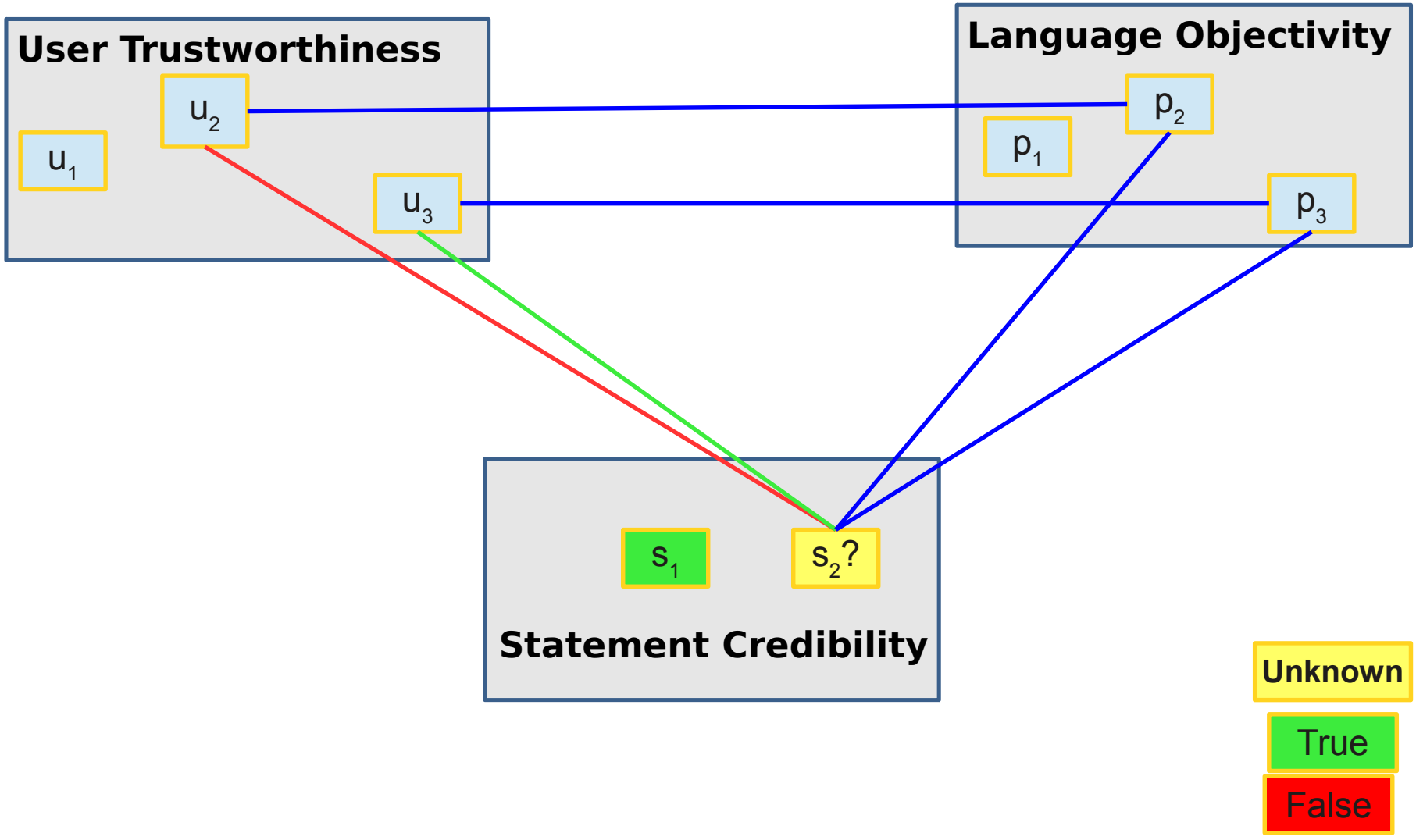


1. Estimate user trustworthiness : $t_k = \frac{\sum_i \mathbb{1}_{S_{i,k}=\text{True}}}{|S_k|}$



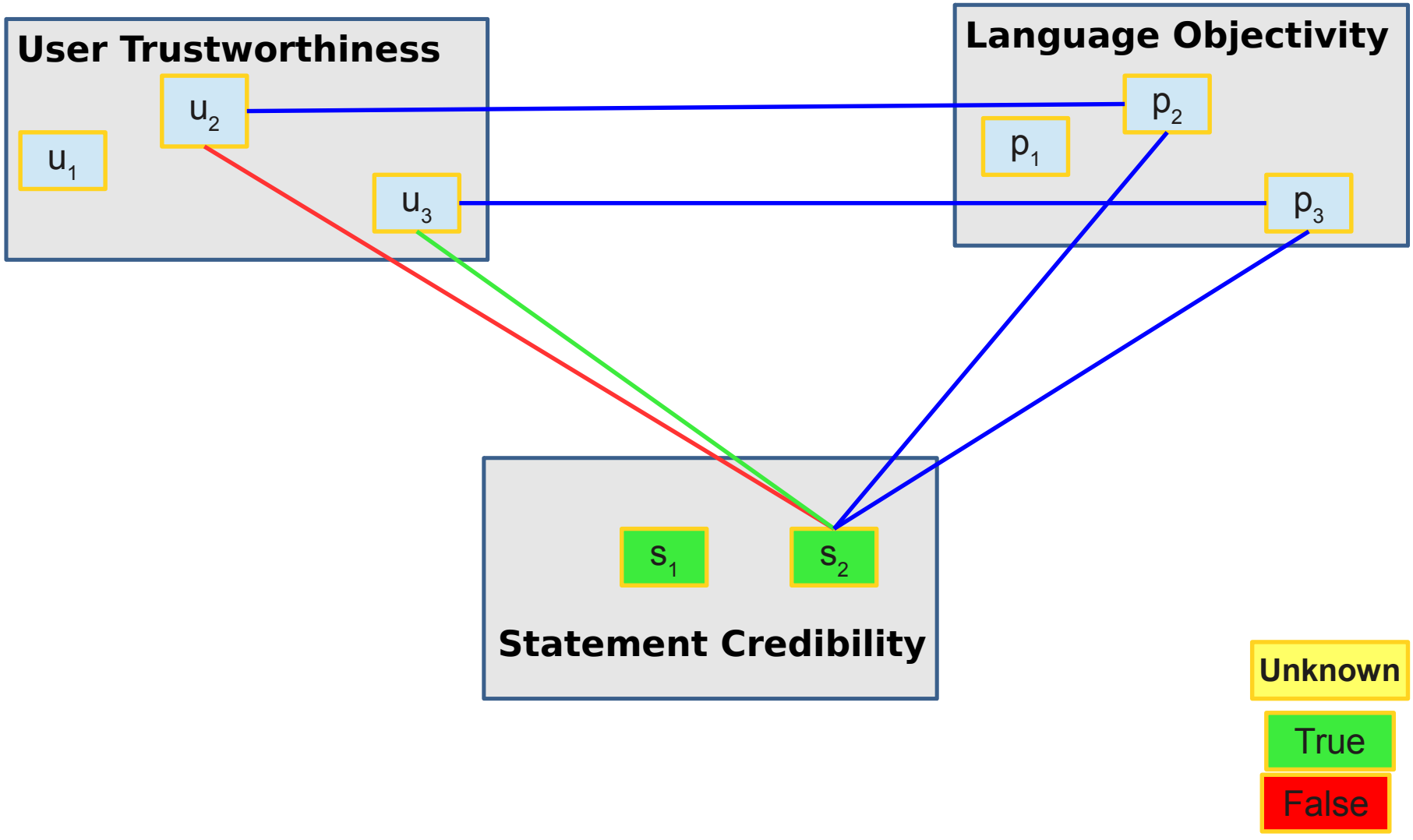
2. E-Step : Estimate label of unknown statements by Gibbs' sampling :

$$Pr(S_i^U | P, U, S^L; W) \propto \prod_{\nu \in C} t_k \times \phi_\nu(S_\nu^*, p_j, u_k; W)$$



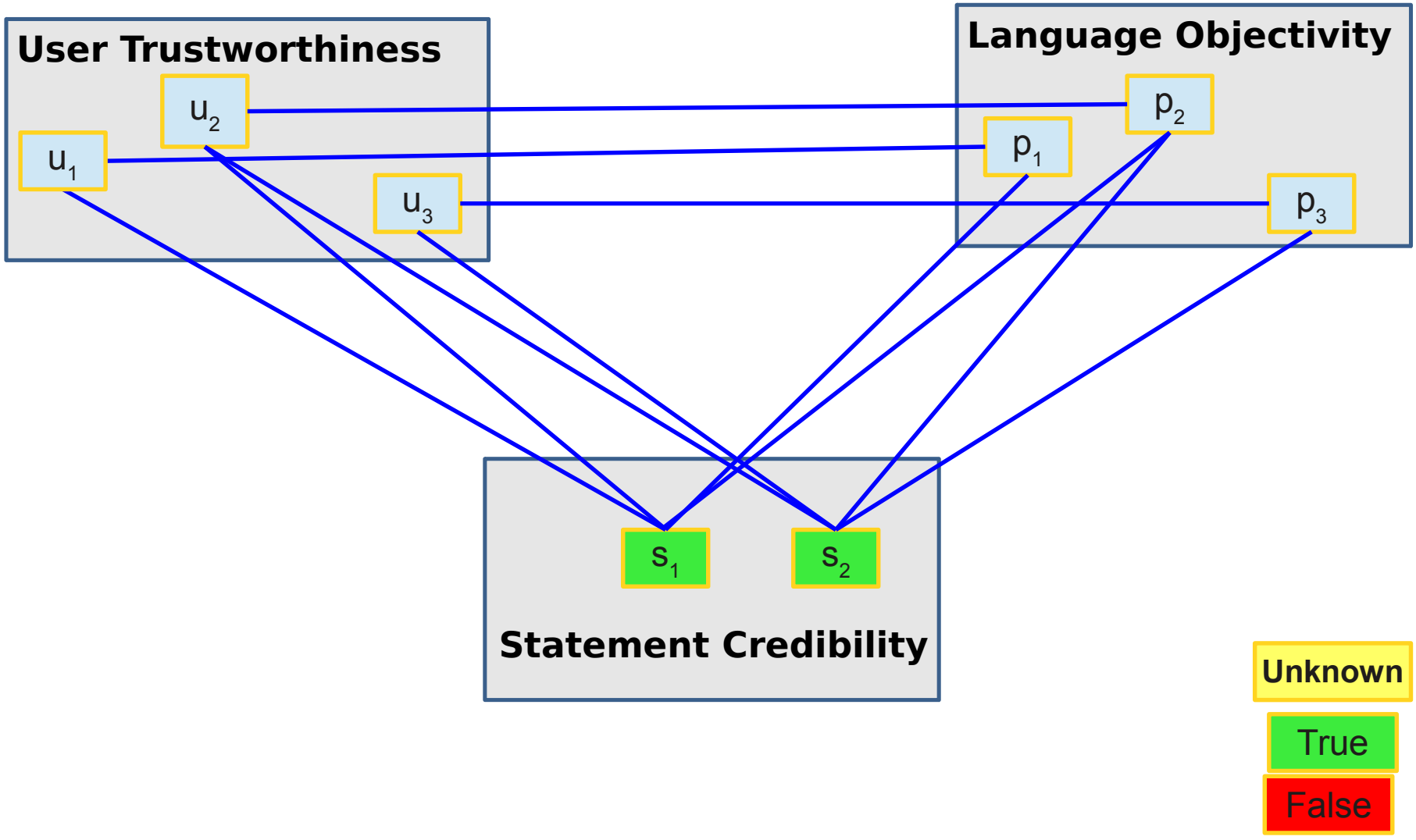
2. E-Step : Estimate label of unknown statements by Gibbs' sampling :

$$Pr(S_i^U | P, U, S^L; W) \propto \prod_{\nu \in C} t_k \times \phi_\nu(S_\nu^*, p_j, u_k; W)$$

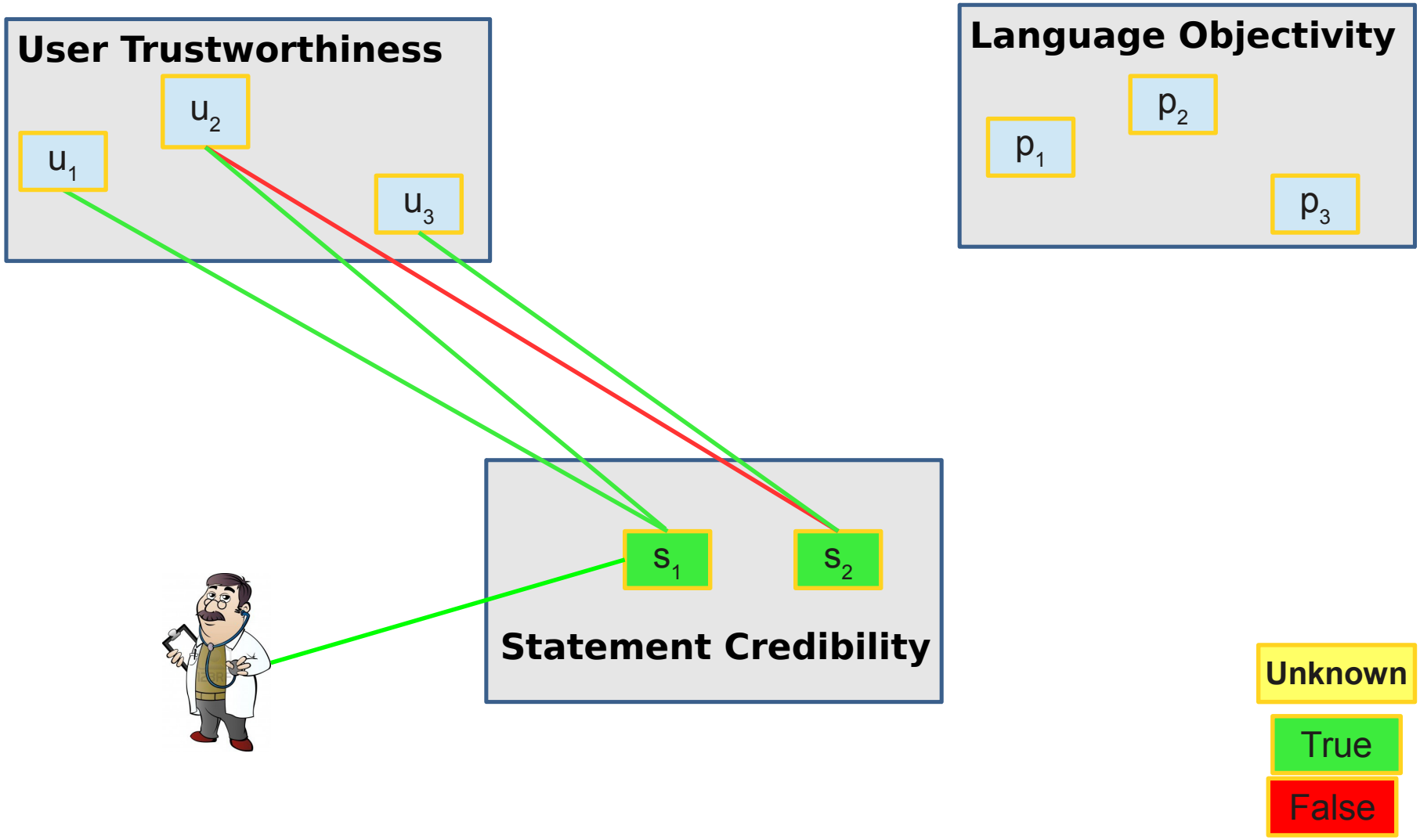


3. M-Step : Maximize log-likelihood to estimate feature weights using Trust Region Newton :

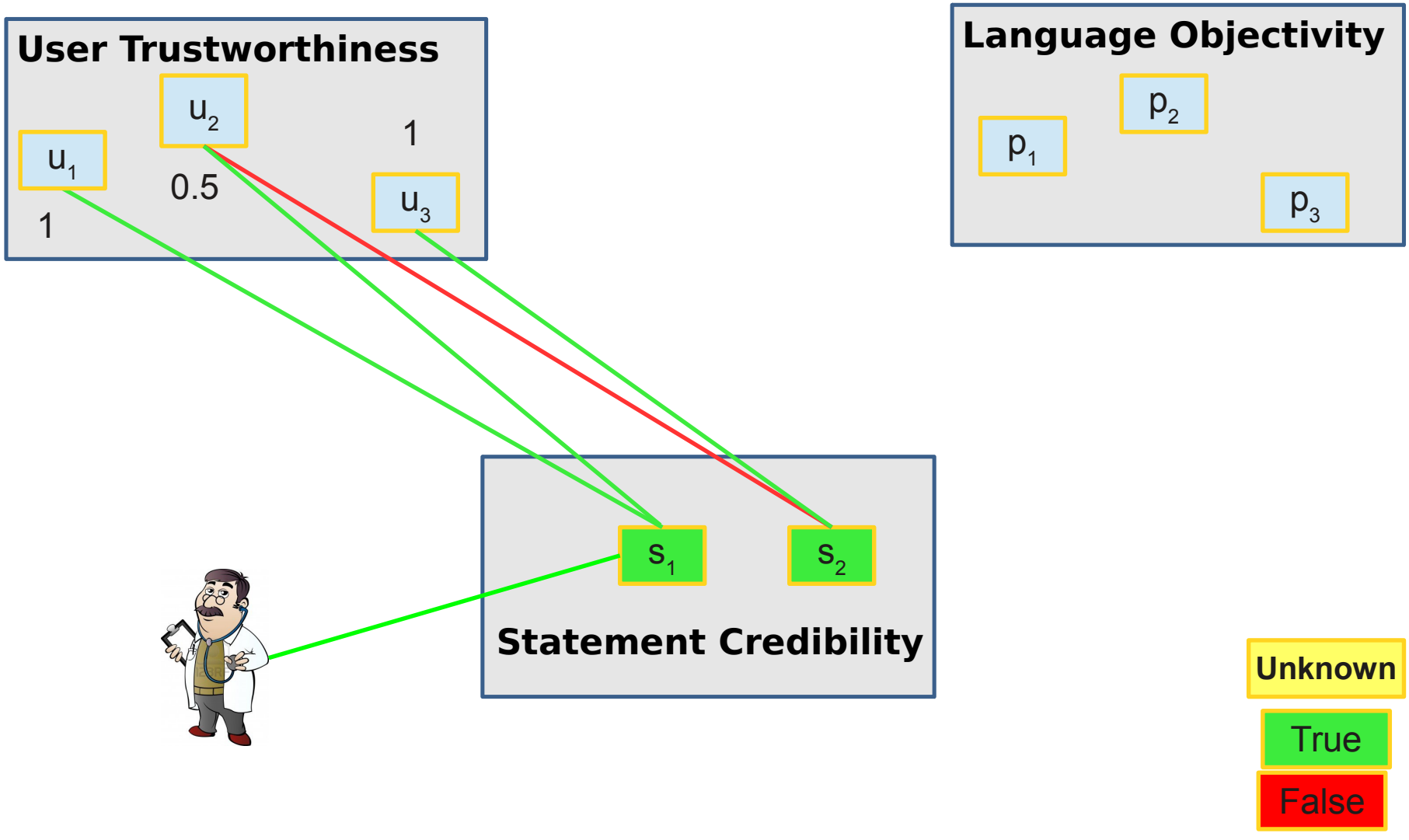
$$W^{(\nu+1)} = argmax_{W'} \sum_{S^U} q(S^U) \log Pr(S^L, S^U | P, U; W')$$



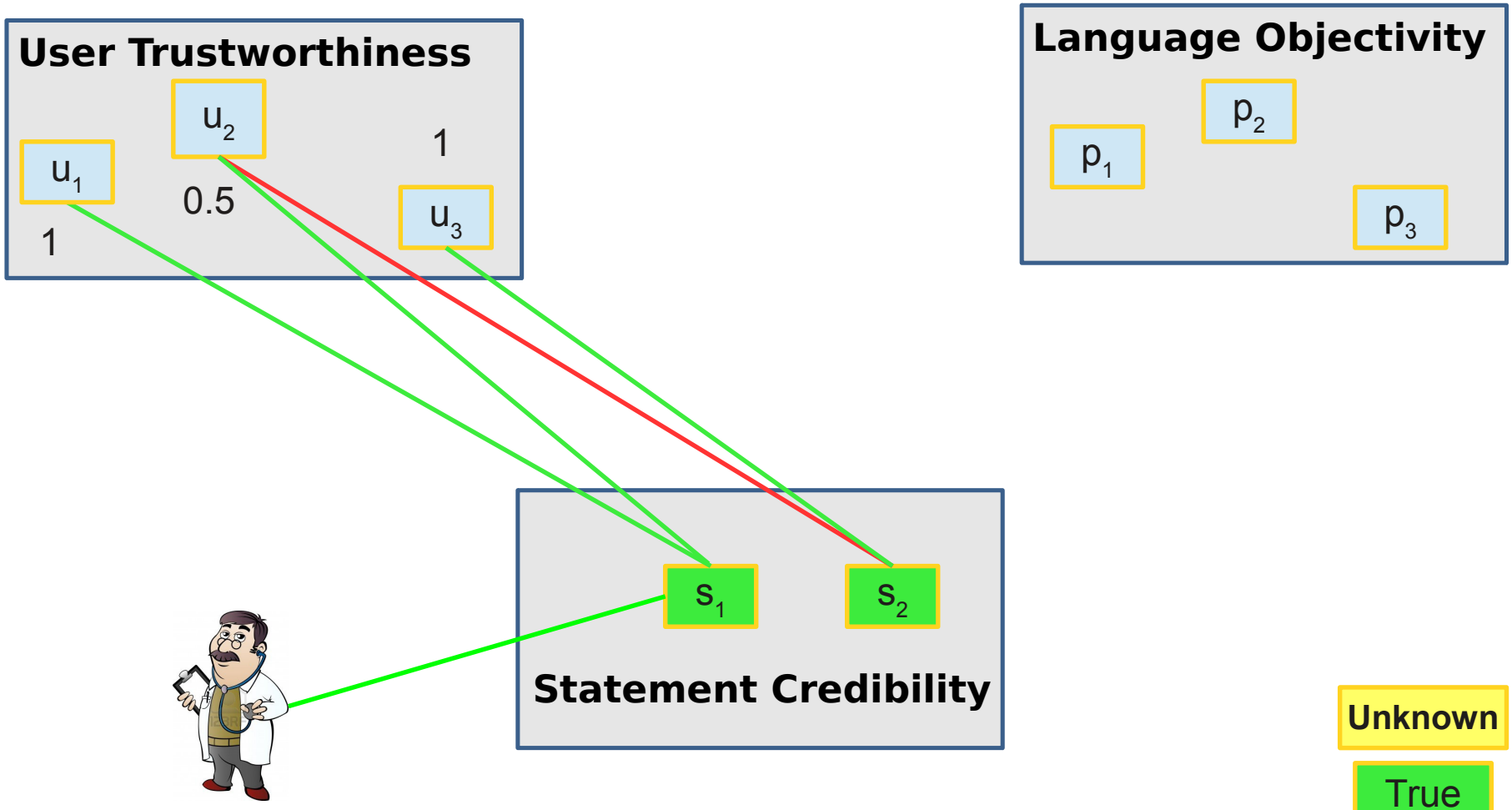
4. Re-Estimate user trustworthiness : $t_k = \frac{\sum_i \mathbb{1}_{S_{i,k}=\text{True}}}{|S_k|}$



4. Re-Estimate user trustworthiness : $t_k = \frac{\sum_i \mathbb{1}_{S_{i,k}=\text{True}}}{|S_k|}$



4. Re-Estimate user trustworthiness : $t_k = \frac{\sum_i \mathbb{1}_{S_{i,k}=\text{True}}}{|S_k|}$



5. Apply *E-Step* and *M-Step* until convergence

Dataset

Healthboards.com community (`www.healthboards.com`) with 850,000 registered users and 4.5 million messages

- We sampled 15,000 users with 2.8 million messages

Expert labels about drugs from MayoClinic portal

- 2172 drugs categorized in 837 drug families
- 6 widely used drugs used for experimentation

Dataset

Healthboards.com community (`www.healthboards.com`) with 850,000 registered users and 4.5 million messages

- We sampled 15,000 users with 2.8 million messages

Expert labels about drugs from MayoClinic portal

- 2172 drugs categorized in 837 drug families
- 6 widely used drugs used for experimentation

Drug Statistics[a]

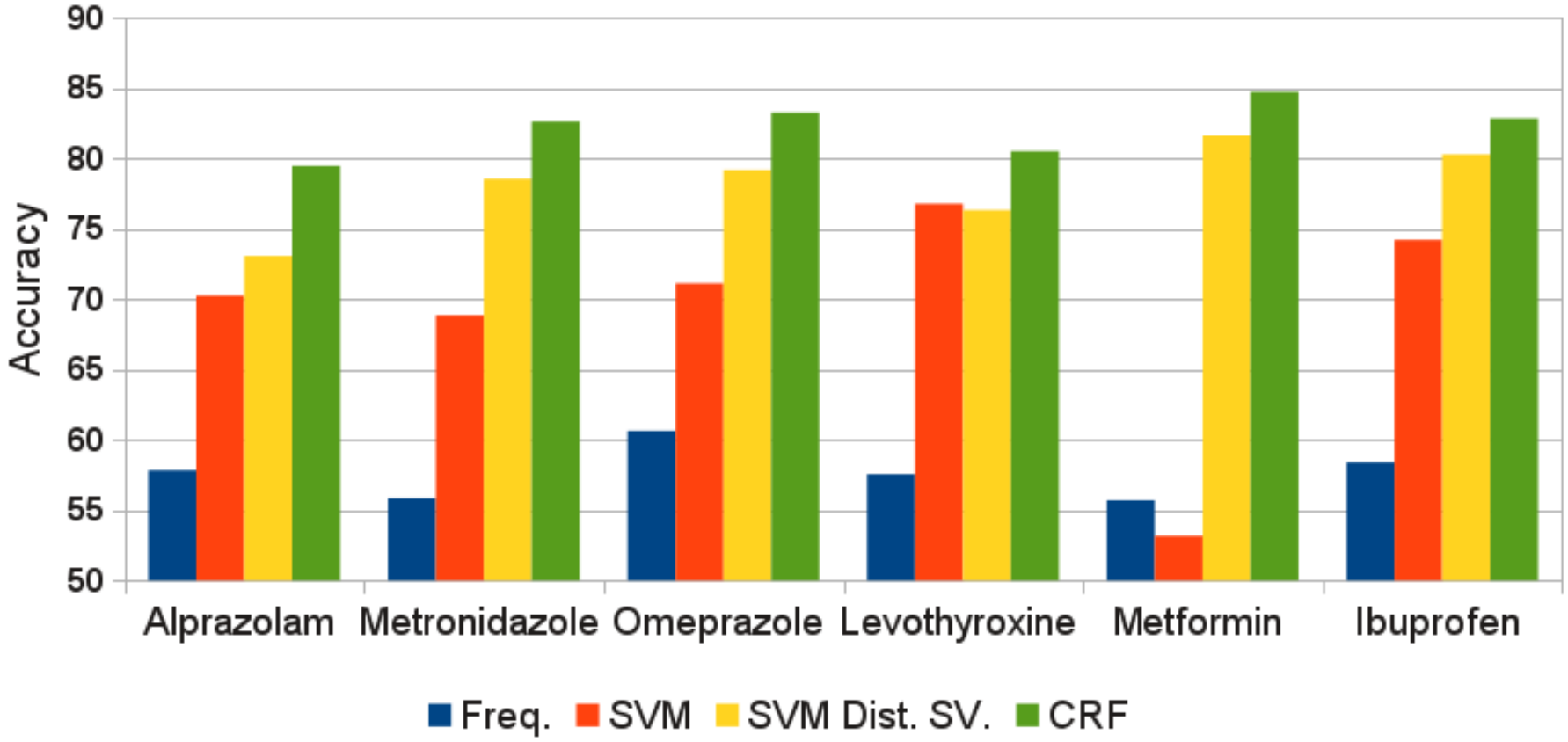
[a]Data available at : http://www.mpi-inf.mpg.de/impact/peopleondrugs/

Drugs	Treatment For	# Users	# Posts
alprazolam	anxiety, depression, panic disorder	2.8K	21K
ibuprofen	pain, symptoms of arthritis	5.7K	15K
omeprazole	acidity in stomach and ulcers	1K	4K
metformin	high blood sugar, diabetes	.8K	3.6K
levothyroxine	hypothyroidism	.4K	2.4K
metronidazole	bacterial infection	.5K	1.6K

Baselines

- Frequency of statements
- SVM Classification
  - Feature vector for each statement using all our features
- SVM Classification with Distant Supervision
  - Each user, post and statement instance constitutes a feature vector
  - Aggregate labels of all such instances for a statement by majority voting

Accuracy Comparison



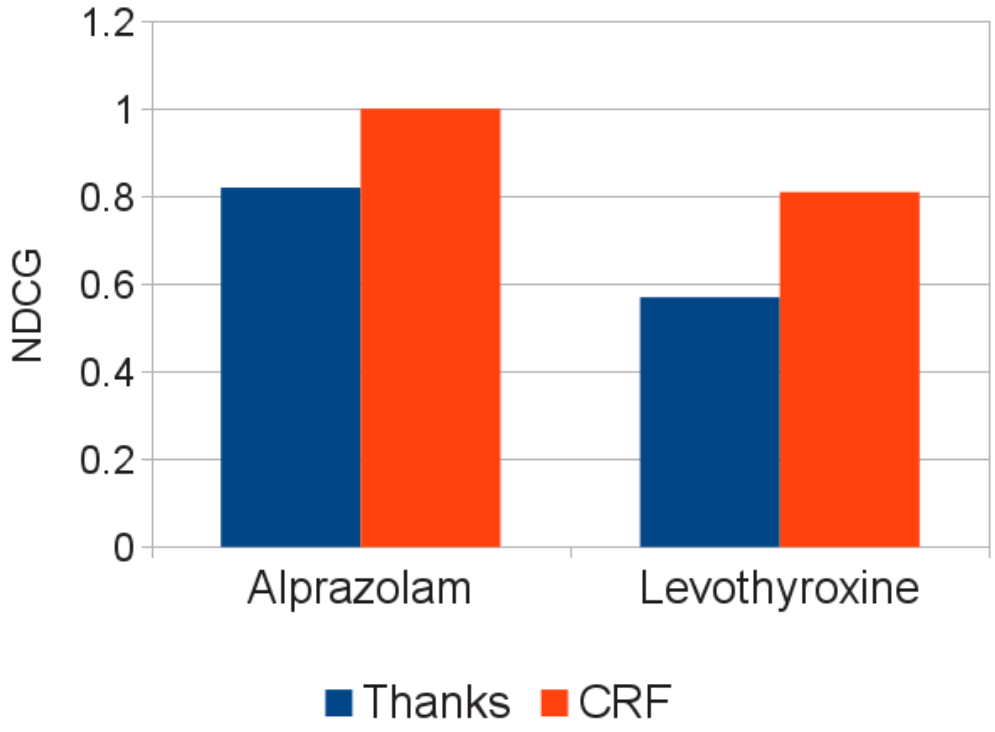
Use-Case: Following Trustworthy Users

What users should I follow to get information on drug *X* ?

Baseline: Rank users based on #*thanks* from community

Use-Case: Following Trustworthy Users

Compare with human annotations



Conclusions

Proposed a probabilistic graphical model to jointly learn *user trustworthiness, statement credibility and language use*

- To extract side-effects of drugs from communities
- Identify expert users

Provides a framework to incorporate richer linguistic (e.g., bias, discourse) and user (e.g., perspective, expertise) features

Thank you