

Jointly Modeling Aspects, Ratings and Sentiments for Movie Recommendation

Chao-Yuan Wu, CMU

Joint work with Qiming Diao¹, Minghui Qiu¹,
Alexander J. Smola^{2,3}, Jing Jiang¹, Chong Wang²

¹SMU, ²CMU, ³Google

Movie review



a real thriller with a good mix of older 007 and new



Daniel Craig plays a believable James Bond. Eve plays well with grit and feminism including a surprise revelation at the end. It's touching as well with the final scenes in the mansion and the old Caretaker. Adele's award for best song is well deserved. But the plot was pretty weak and the film dragged on and on and on, probably being 30 minutes too long.

Movie review



a real thriller with a good mix of older 007 and new



Daniel Craig plays a believable James Bond. Eve plays well with grit and feminism including a surprise revelation at the end. It's touching as well with the final scenes in the mansion and the old Caretaker. Adele's award for best song is well deserved. But the plot was pretty weak and the film dragged on and on and on, probably being 30 minutes too long.

Movie review

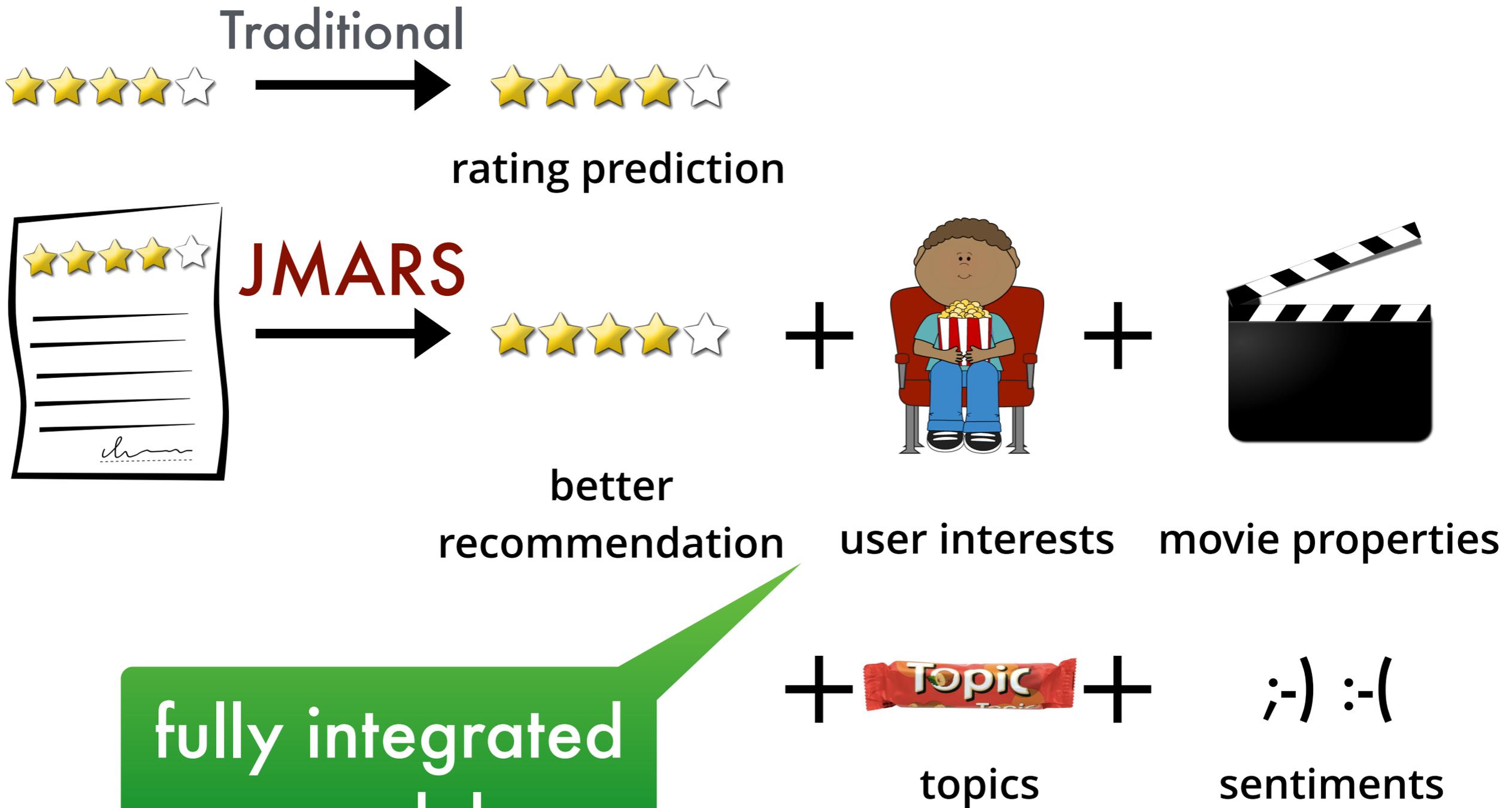


a real thriller with a good mix of older 007 and new



Daniel Craig plays a believable James Bond. Eve plays well with grit and feminism including a surprise revelation at the end. It's touching as well with the final scenes in the mansion and the old Caretaker. Adele's award for best song is well deserved. But the plot was pretty weak and the film dragged on and on and on, probably being 30 minutes too long.

Motivation



fully integrated model

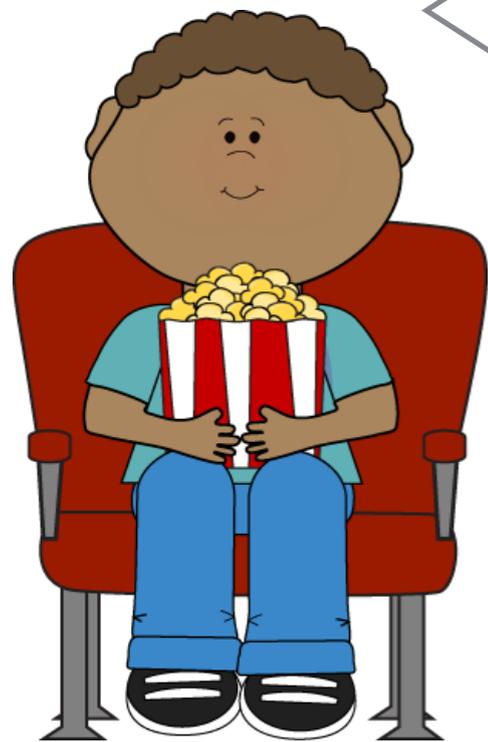
Aspect sentiment modeling



Aspect	Rating	Dist.
Acting	2.0	
Story	7.5	
Horror	-	
Romance	6.0	
...		



Aspect sentiment modeling



Aspect

Rating

Dist.

?

?

?

?

?

?

?

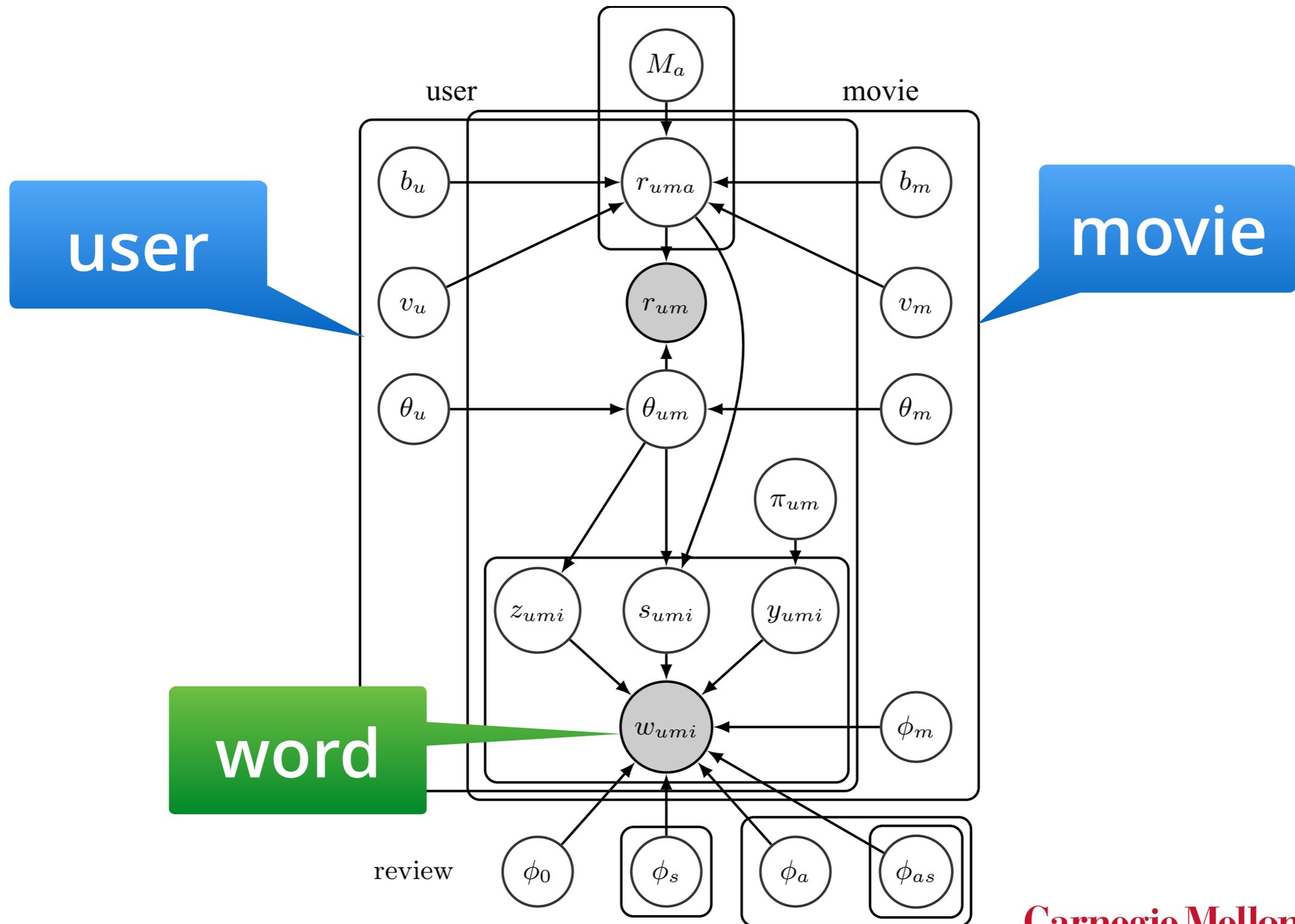
?

?

?



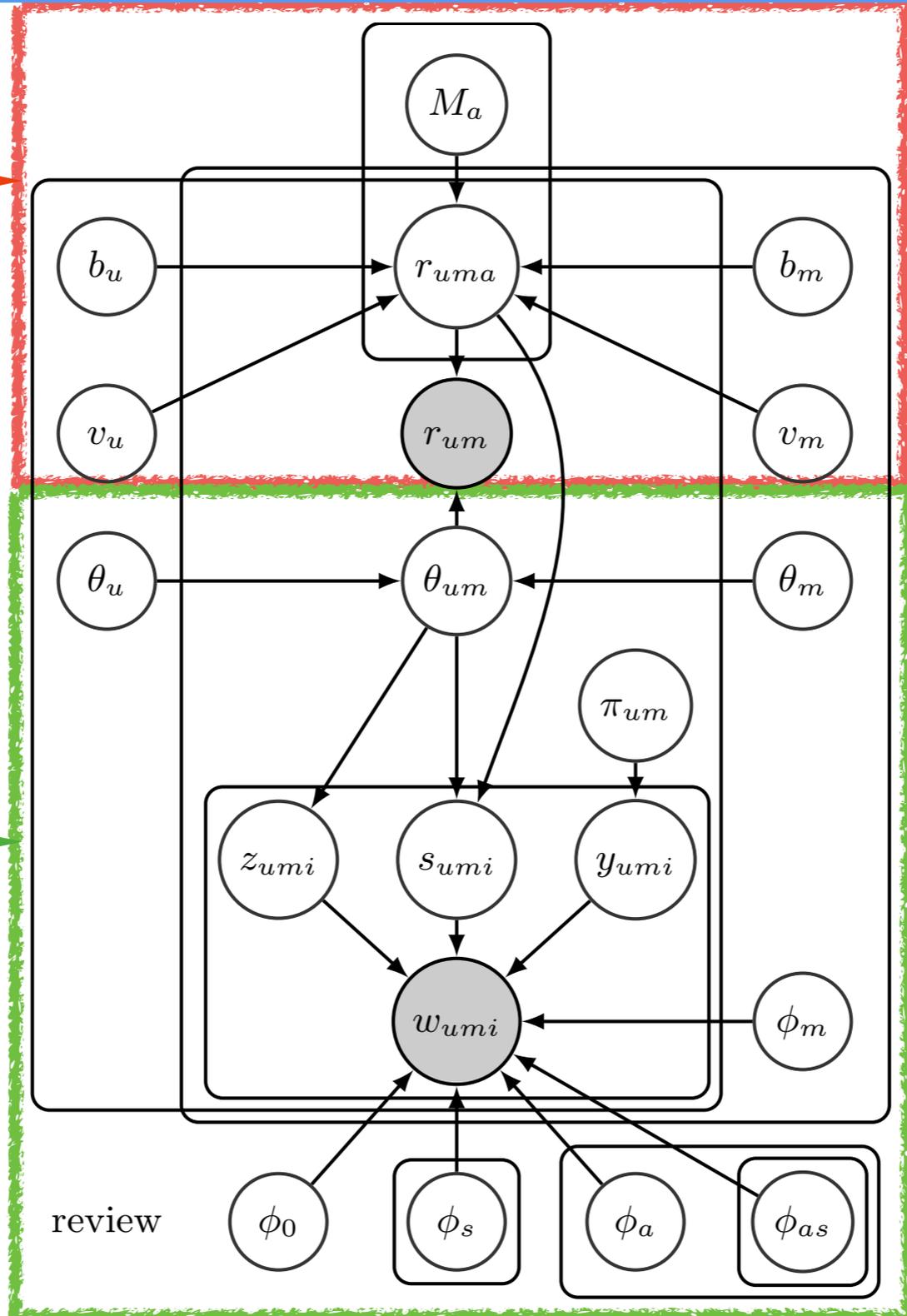
Everything but the kitchen sink ...



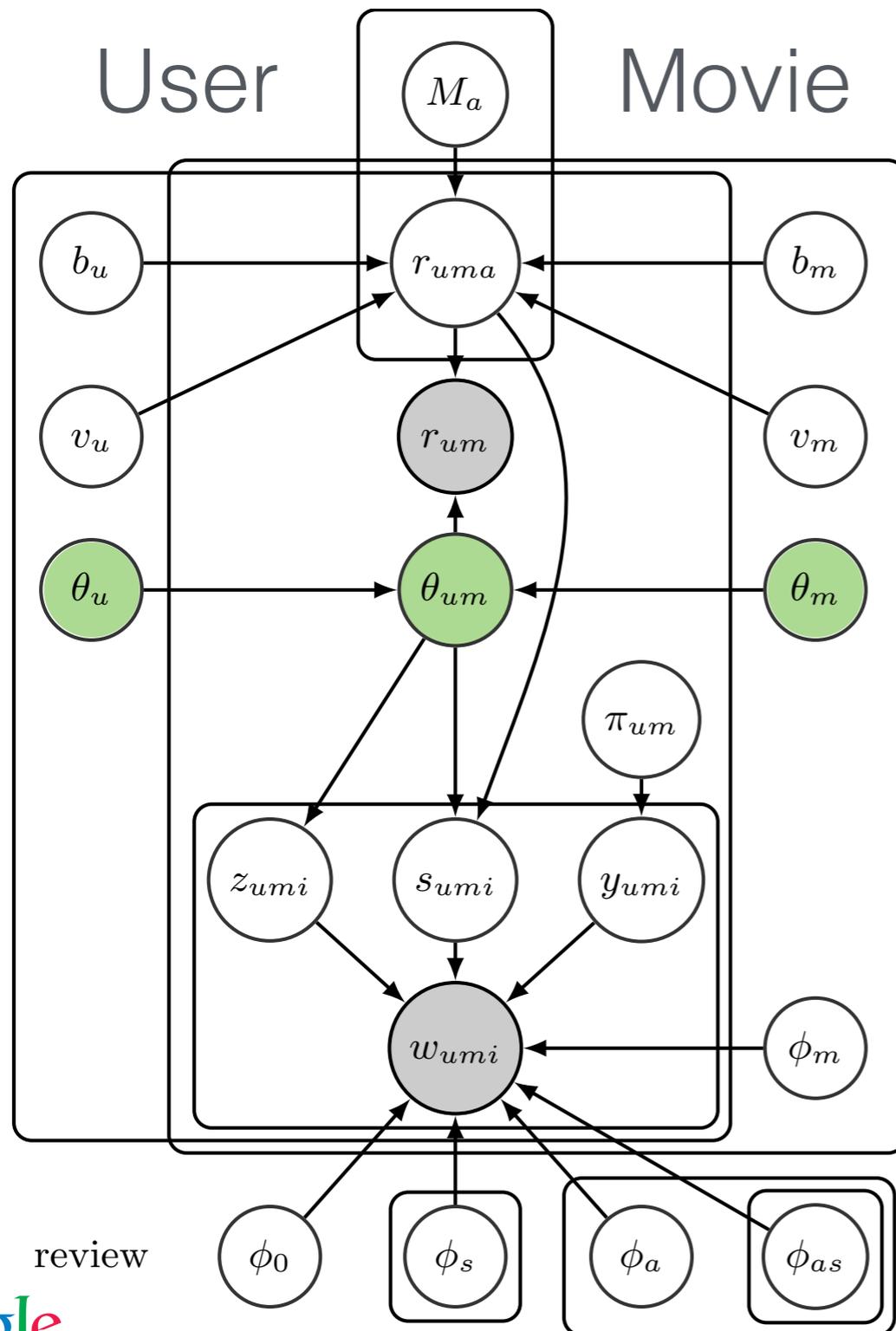
Everything but the kitchen sink ...

rating
model

review
model



Aspect modeling

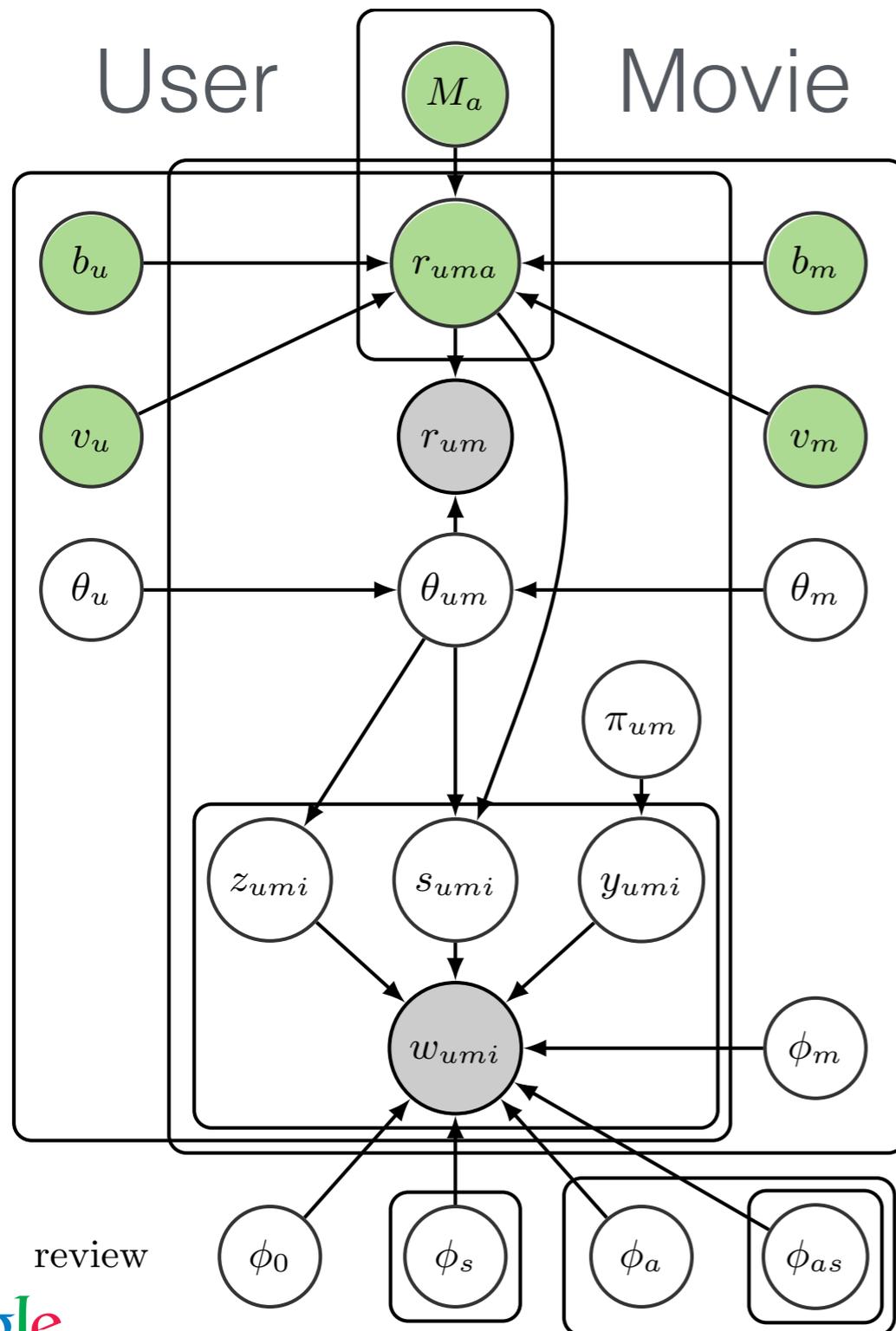


$$\theta_u \sim \mathcal{N}(0, \sigma_{\text{user aspect}}^2 \mathbf{I})$$

$$\theta_m \sim \mathcal{N}(0, \sigma_{\text{movie aspect}}^2 \mathbf{I})$$

$$\theta_{um} \propto \exp(\theta_u + \theta_m)$$

Rating modeling



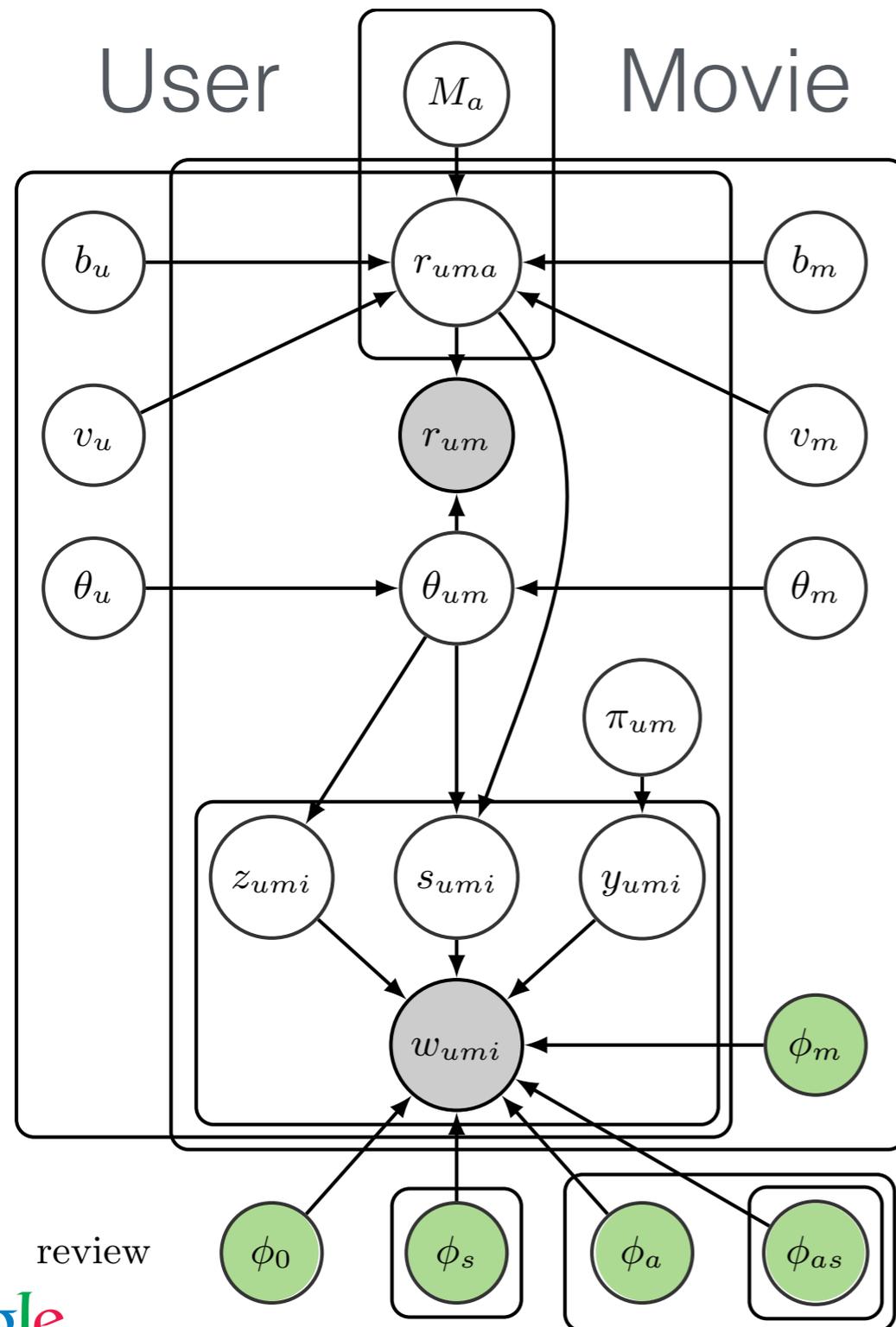
Aspect-specific rating

$$r_{uma} = v_u^\top M_a v_m + b_u + b_m$$

$$\hat{r}_{um} = \sum_a \theta_{uma} r_{uma}$$

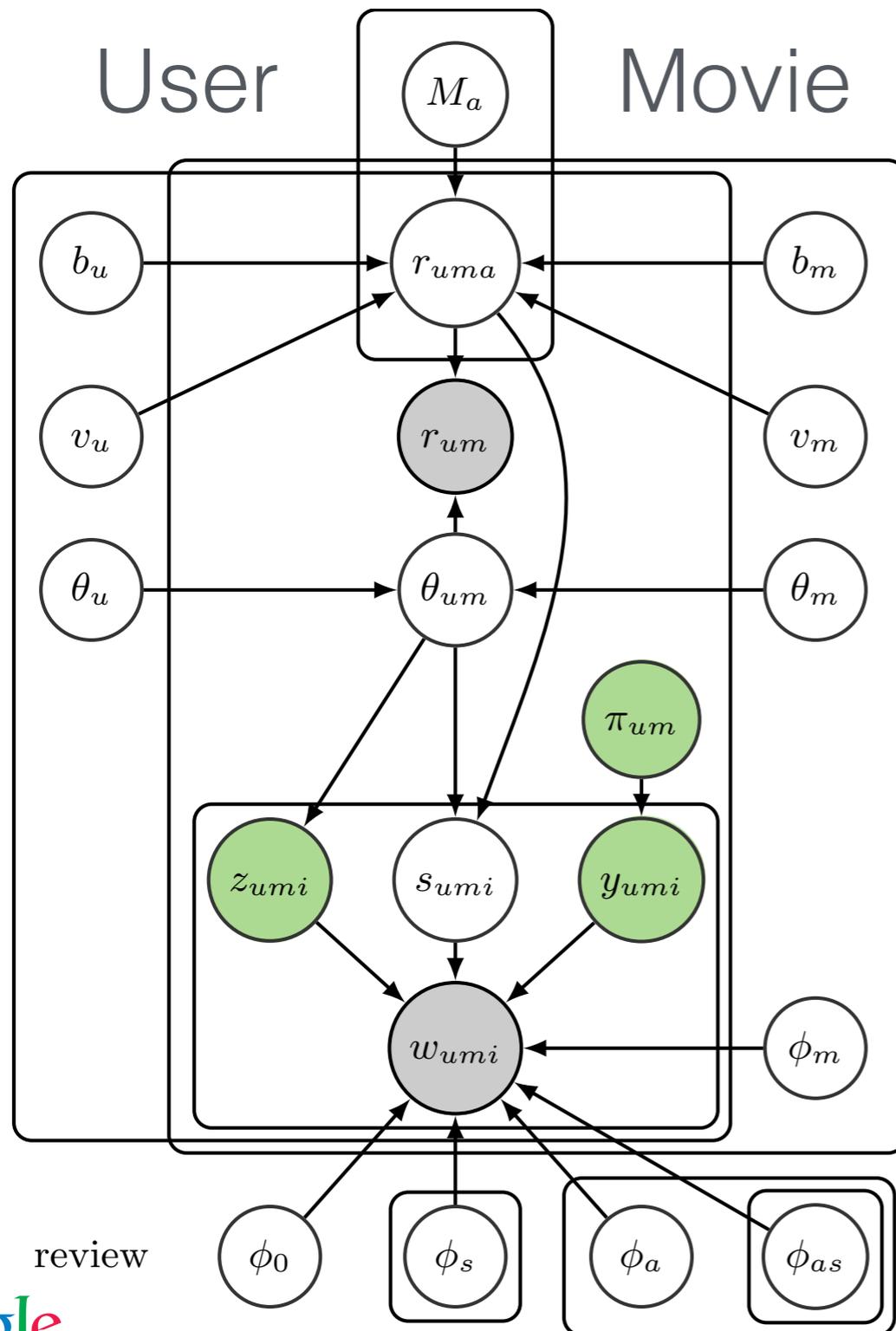
Overall rating

Language model



- Background
- Sentiment
- Movie-specific
- Aspect-specific
- Aspect-specific sentiment

Topic model

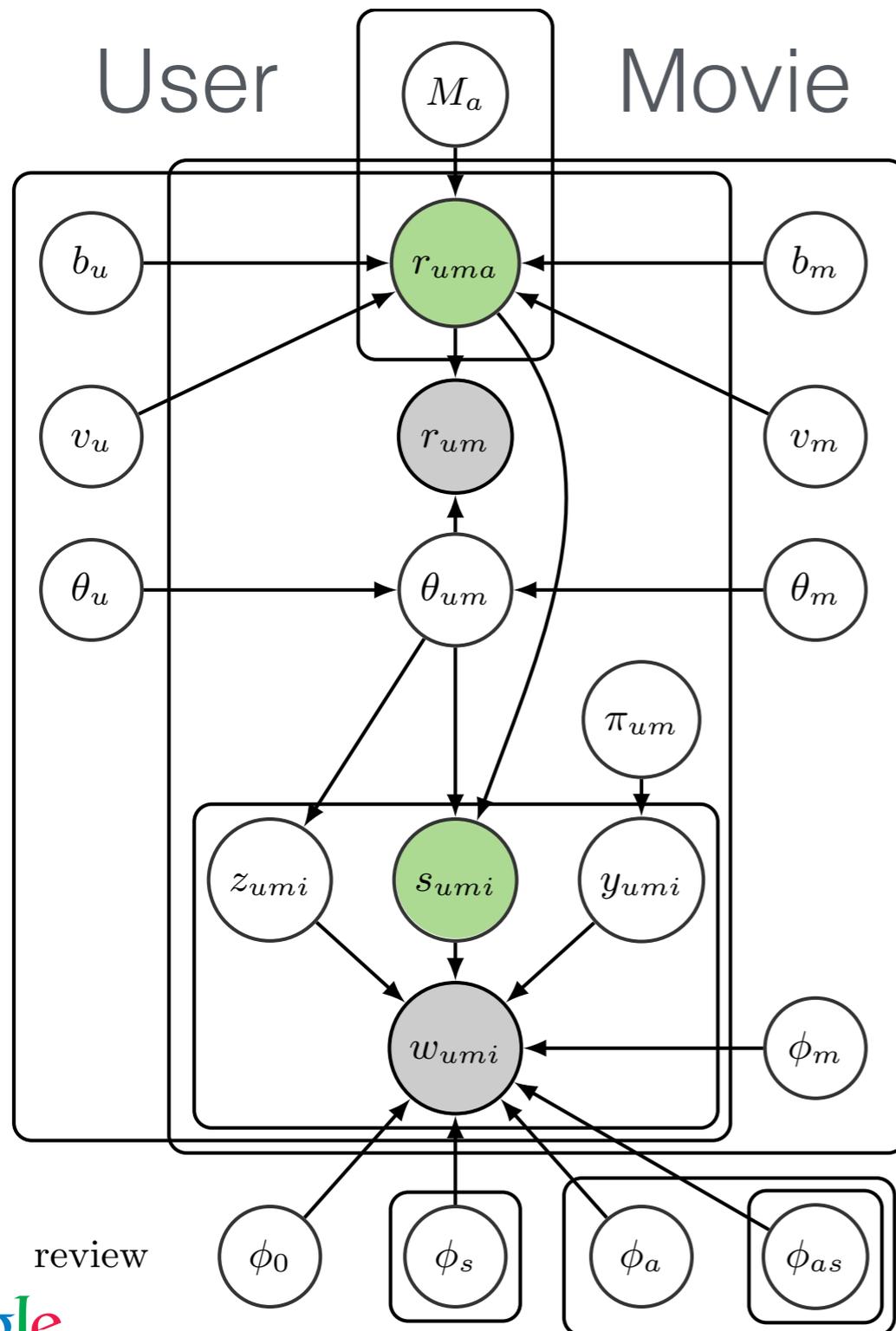


$$\pi_{um} \sim \text{Dir}(\gamma) \text{ for } \pi_{um} \in \mathcal{P}_5$$

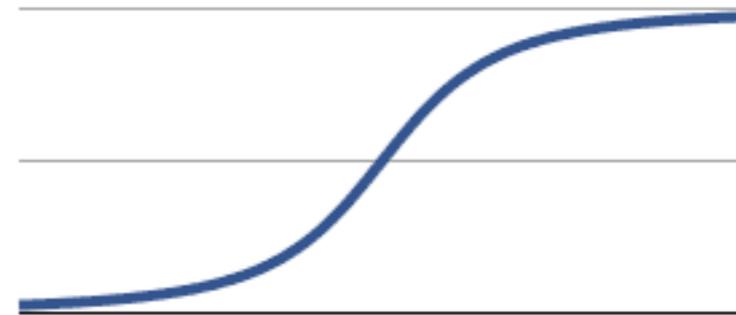
$$y_{umi} \sim \text{Mult}(\pi_{um})$$

$$z_{umi} \sim \text{Mult}(\theta_{um})$$

Sentiment Mapping



Sigmoid mapping



$$p(s_{umi} | r_{uma}, z_{umi} = a)$$

$$= \frac{1}{1 + e^{-s_{umi}(r_{uma} - b)}}$$

Aspect-specific rating

Inference by hybrid sampling / optimization

Objective $\mathcal{L} \text{ “=”} \sum_{r_{um} \in \mathcal{R}} [\epsilon^{-2} (r_{um} - \hat{r}_{um})^2 - \log p(w_{um} | \Upsilon, \Omega)]$

E step: fix $\Theta = [\{v_u, b_u, \theta_u\}_{u=1}^U, \{v_m, b_m, \theta_m\}_{m=1}^M, \{M_a\}_{a=1}^A]$
sample on $\{y, z, s\}_{umi}$

M step: fix $\{y, z, s\}_{umi}$
optimize Θ

see e.g. Wallach et al. 2006



Experiments

Instantly watch 40,000
movies & TV episodes

amazon
Prime instant video

IMDb

Find Movies, TV shows, Celebrities and more...

All



IMDbPro

IMDb Apps

Help

Movies, TV
& Showtimes

Celebs, Events
& Photos

News &
Community

Watchlist

Login



The Young Ones
HD Teaser Trailer

See featured HD Trailers »

54k users
22k movies
348k reviews
118k unigrams (tokens)
0.03% density (Netflix has 1%)

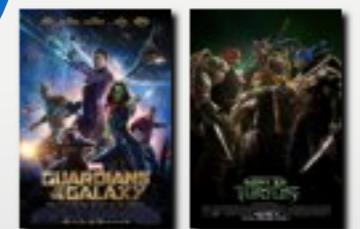
Emmy Awards This Monday!



The 66th Primetime Emmy Awards are this Monday, airing on CBS at 8pm ET/5pm PT, hosted by Seth Meyers. Get ready for the show with our Road to the Emmys section, featuring the full list of major nominees, photos from last week's Creative Arts Emmys, all the glamour from last year's

playing near you

See all showtimes »



Opening This Week

Sin City: A Dame to Kill For [Get Tickets](#)

When the Game Stands Tall [Get Tickets](#)

If I Stay [Get Tickets](#)

Mean squared error (MSE)

$$\text{MSE} = \frac{1}{N_{\text{test}}} \sum_{r_{um} \in \mathcal{R}_{\text{test}}} (r_{um} - \hat{r}_{um})^2$$

Factor size	Offset	PMF ¹	HFT ²	Ours ³		
				A=20	A=10	A=5
5	7.07	5.99	5.21	4.97	5.11	5.23
10		5.92	5.14	5.05	5.18	5.28

PMF: Probabilistic matrix factorization [Mnih & Salakhutdinov 2007].

HFT: Hidden Factors as Topics [McAuley & Leskovec 2013].

A = number of aspects.

Perplexity

$$\log \text{PPX}(D_{\text{test}}) = -\frac{1}{N_w} \sum_{\substack{u,m \\ d_{um} \in D_{\text{test}}}} \sum_i \log p(w_{umi})$$

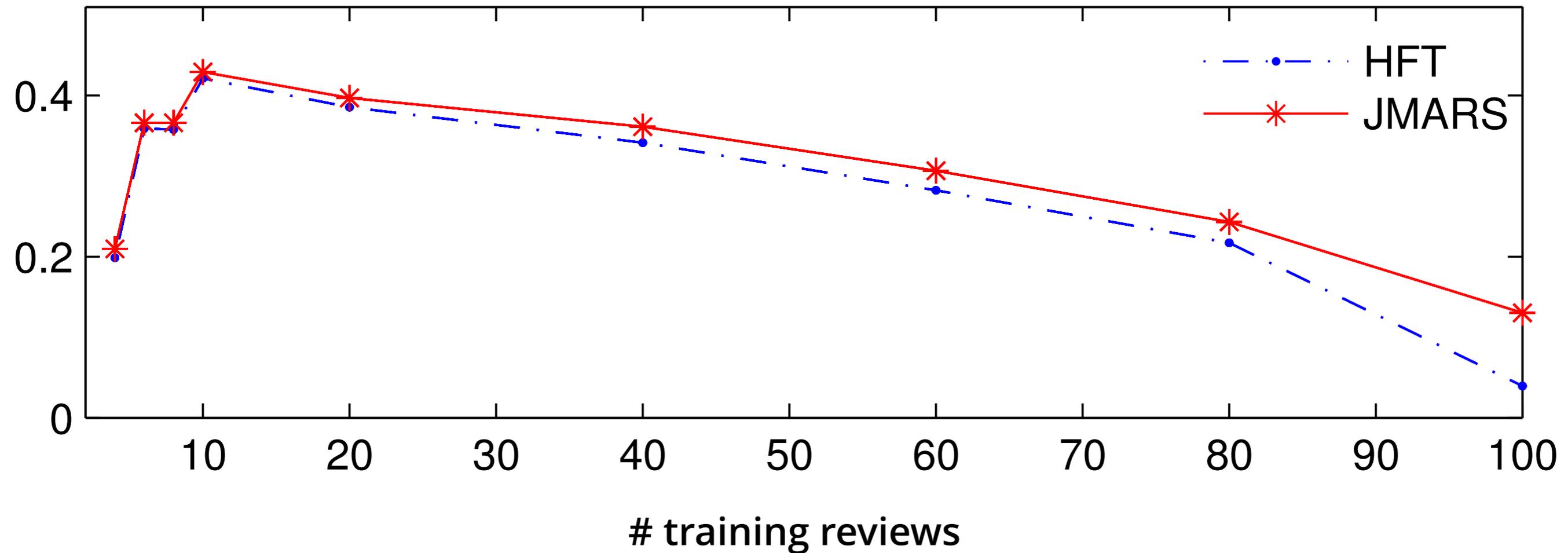
Factor size	HFT	Ours		
		A=20	A=10	A=5
5	8,247	3,348	3,369	3,399
10	7,459	3,335	3,359	3,379

HFT: Hidden Factors as Topics [McAuley & Leskovec 2013].

A = number of aspects.

Cold start

Improvement in MSE compared to PMF for 'cold-start' users.



Qualitative evaluation

I enjoyed this DVD from the library very much. Daniel Craig plays a believable James Bond. There are some of the older 007 action scenes and similar gimmicks with updates thanks to the younger Quartermaster. Eve plays well with grit and feminism including a surprise revelation at the end. It's touching as well with the final scenes in the mansion and the old Caretaker. Adele's award for best song is well deserved. But the plot was pretty weak and the film dragged on and on and on, probably being 30 minutes too long. The filming is it's usual high quality, but still overall both my wife and I found this boring, something you can't usually level against a Bond film.

Sentiment & movie-specific words

Aspect	Director	History	War	Life	Character
Rating	9.36	8.55	8.51	9.20	9.50
Prob	0.12	0.10	0.09	0.09	0.08

...what an **excellent** piece of cinema ...the actors are great and directing **incredible** ...in **300**, **Gerard Butler** **dominates** the screen ...battle scenes are **incredible** ...

Positive words

Movie-specific

Aspect rating

Aspect	Director	History	War	Life	Character
Rating	9.36	8.55	8.51	9.20	9.50
Prob	0.12	0.10	0.09	0.09	0.08

...what an excellent piece of cinema ...the actors are great and **directing incredible** ...in 300, Gerard Butler **dominates the screen** ...**battle scenes are incredible** ...

Aspect & Aspect-sent words

violence			human/social		
Aspect	Negative	Positive	Aspect	Negative	Positive
murder	nudity	violence	social	attempts	ultimately
police	exploitation	cinematography	moral	result	level
killer	female	genre	society	sense	dramatic
crime	pace	solid	point	material	intelligent
cop	violent	gritty	question	superficial	contemporary
mystery	tension	stylish	human	shallow	strength
detective	nasty	highly	god	taut	essentially
case	gratuitous	sinister	act	trite	complexity
death	naked	inspired	nature	ugly	sympathetic
dead	sleazy	engaging	issues	one-dimensional	genuinely
thriller	brutal	macabre	men	grotesque	compelling
victim	hilarious	blood	personal	banal	understated
murdered	murderous	vicious	culture	awkward	sharp
murders	low-budget	nasty	behavior	satirical	equally
criminal	trashy	compelling	conflict	excessive	thoughtful

Movie specific

Movie	Words
Casino Royale	bond, craig, james, casino, royale
Batman Begins	batman, bruce, wayne, bale, begins
The Matrix Reloaded	matrix, neo, reloaded, action, flight
American Beauty	beauty, american, spacey, lester, kevin

Background tokens

Type	Words
Background	film, story, character, films, characters, movies, scenes, scene, real, good
Positive	great, good, love, acting, fun, funny, excellent, lot, thought, perfect
Negative	bad, good, pretty, acting, plot, boring, worst, watching, minutes, stupid

Summary

- **Joint model** of ratings, reviews, movies, users
 - **No aspect annotation needed**
 - **Automatic topic generation**
 - **Automatic sentiment extraction**
 - **Identify *why* users like movies**
 - **Movie specific and user specific aspects**
- **Everyone else needs aspects & sentiment labels**