Learning Through the Grapevine

Matthew O. Jackson (Stanford)
Suraj Malladi (Cornell)
David McAdams (Duke)

This talk

Social learning

Some hear primary info

Relay messages with noise

Q: How hard to learn as distance to primary source grows?

A: Typically impossible

Examples of noise

Simmons, Adamic, Adar (2011) account of tweets:

“Street style shooting in Oxford Circus for ASOS and Diet Coke. Let me know if you're around!”

became

“Shooting in progress in Oxford Circus? What?"

became

“Shooting in progress in Oxford Circus, stay safe people.”

within 3 minutes...

Examples of noise

Yellowstone Volcano Observatory (YVO):

Wrote article on repairs for monitoring station damaged during a storm

Tabloid wrote article implying Yellowstone was unmonitored

- Bad copy paste: ‘borehole’ became ‘boreal’

Misspelling and misinformation persisted in social media reports of the incident

Frequency and scope

Adamic et al (2016):

Message about health care reposted ~470,000 times, with ~11% mutation rate and ~100,000 variants

- 121 of the 123 most viral text memes had > 100,000 variants
- Out of 1000s of memes, typical distance in dozens

Gurry (2016)

~500,000 textual variants of Greek New Testament, not including spelling errors

Model of communication

State ω: either 1 or 0

probability θ that $\omega = 1$

Sources hear iid binary signals (1 or 0) about state

Noisily relay signal to neighbors

Mutations

Probability μ_{01} , $\mu_{10} < 1/2$ of message mutating in transmission

μ_{ij} is the probability of an i message becoming a j message

Mutations are iid at each step

Transmission failures

Probability p_1 (p_0) of passing 1 (0)

Message 1 (0) dropped with probability $1 - p_1$ ($1 - p_0$)

E.g.,

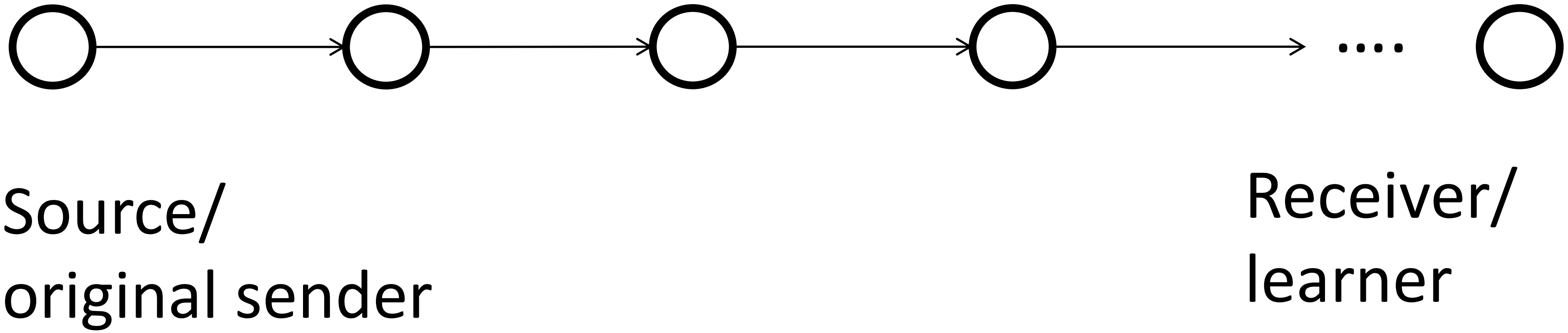
report `news' – significant results with higher prob.

Tell funnier jokes,

Pass salacious rumors…

For now $p_1 = p_0 = p$

Passing along a chain



Multiple chains

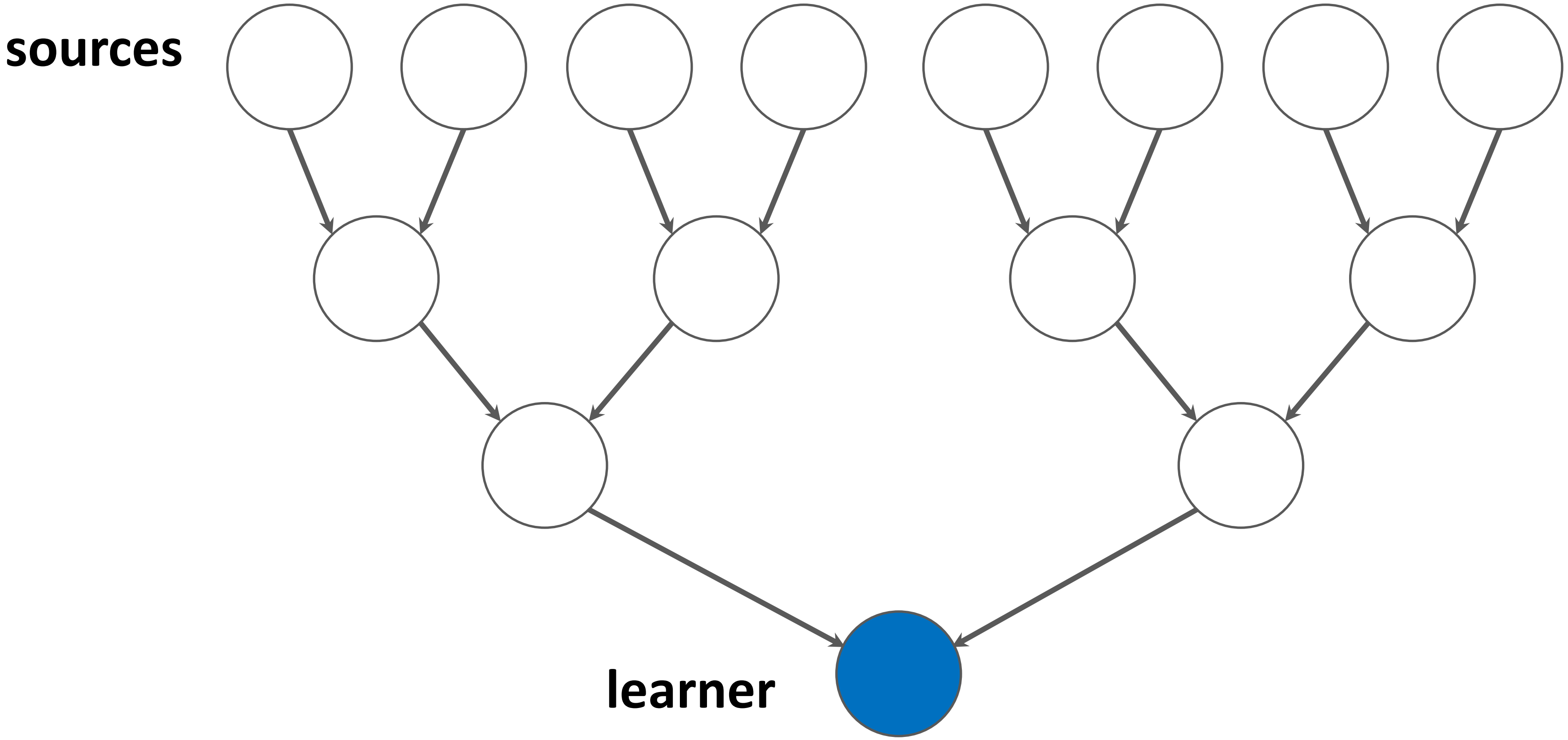
$n(t)$ chains length t, independent passing on each chain

Race: each chain less informative, but more chains

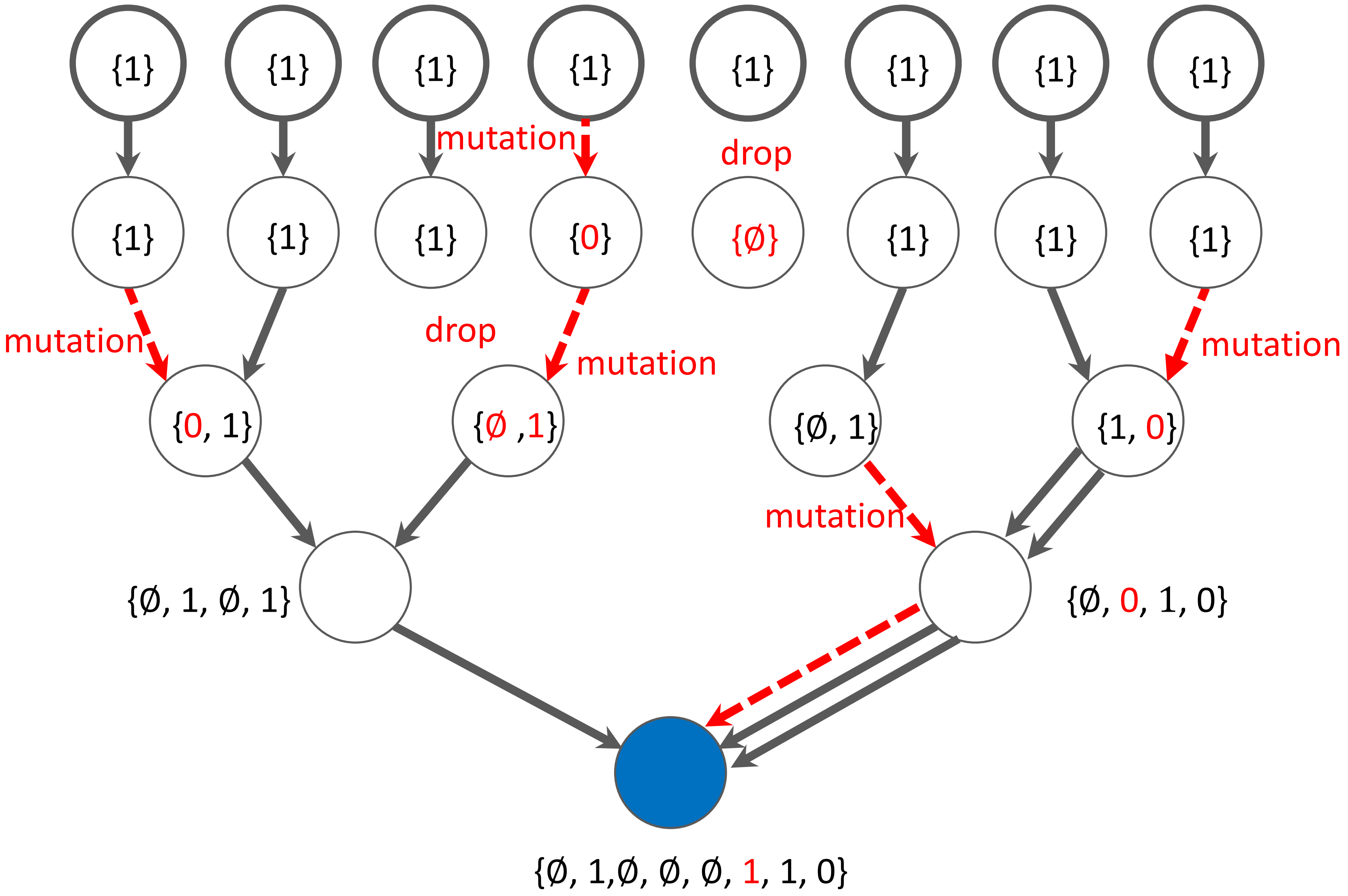
$b(t) = P(\omega = 1 \mid \text{information from } n(t) \text{ iid chains})$

Multiple chains

$t = 3$
$n(t) = 8$



Multiple chains



Learning

$f(t)$ is a threshold for learning if

Plim $b(t) = 1$ or 0 if $n(t)/f(t) \to \infty$

Plim $b(t) = \theta$ if $n(t)/f(t) \to 0$

Note: if $b(t) = 1$ or 0 then beliefs are correct

Exponential threshold for learning

Lemma 1: If $p > 0$ and $\mu_{10}, \mu_{01} < 1/2$, then a threshold for learning is

$$\left(\frac{1}{p(1 - \mu_{10} - \mu_{01})^2}\right)^t.$$

Note: for random tree, $E[\text{degree}] \lessgtr \frac{1}{p\,(1 - \mu_{10} - \mu_{01})^2}$

Impossibility of learning

Proposition 1: If the learner does not know μ_{10} / μ_{01} but has an atomless prior with connected support, then for any $n(t)$, the learner learns nothing in the limit, i.e., Plim $b(t) \; = \; \theta$.

Note: no learning even in case of near-certainty

Finite *t* intuition

$p = 1$

$\mu_{01} = \alpha\pi$

$\mu_{10} = \alpha(1 - \pi)$

α known

π unknown

π-prior supported on $(\frac{1}{3} - \epsilon, \frac{1}{3} + \epsilon)$

Finite *t* intuition

Case 1: Primary sources are direct friends and $n(1)$ large

If $\omega = 1$:

$$\%1s \rightarrow 1 - \mu_{10} > (1-\alpha) + \left(\frac{1}{3} - \epsilon\right)\alpha \equiv L(1)$$

If $\omega = 0$:

$$\%1s \rightarrow \mu_{01} < \left(\frac{1}{3} + \epsilon\right)\alpha \equiv M(1)$$

For small α, $M(1) \ll L(1)$

ω is identified from $\%1s$

i.e., learn if $n(1)$ large

Finite t intuition

Case 2: Primary sources are distant t away and $n(t)$ large

If $\omega = 1$:

$$\%1\text{s} > (1-\alpha)^t + \left(\frac{1}{3} - \epsilon\right)(1-(1-\alpha)^t) \equiv L(t)$$

If $\omega = 0$:

$$\%1\text{s} < \left(\frac{1}{3} + \epsilon\right)(1-(1-\alpha)^t) \equiv M(t)$$

For large t, $M(t) > L(t)$

Even for large $n(t)$, can rationalize either ω for %1s $\in (L(t), M(t))$

Learn partially from the likelihood of rationalizing π's

Limit t intuition

Probability that $\%1s \in (L(t), M(t))$ goes to 1

Rationalizing π's converge to each other

No learning

Comment: role of asymmetries

Proposition 1': If $\mu_{10} = \mu_{01} = \mu$ and the learner does not know μ and has a prior bounded away from 1/2, then there exists $f(t)$ such that

$$\text{Plim } b(t) = 1 \text{ or } 0 \qquad \text{if } n(t)/f(t) \to \infty.$$

Upshot

When primary sources are distant

If people start out with heterogenous beliefs

Even w/ Bayesian agents

No amount of communication $\Longrightarrow$ consensus

Extensions:

Asymmetric message survival rates

Sources can be near and far

Thanks!