АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

Сергей Николенко

НИУ ВШЭ — Санкт-Петербург
15 сентября 2017 г.

Random facts:

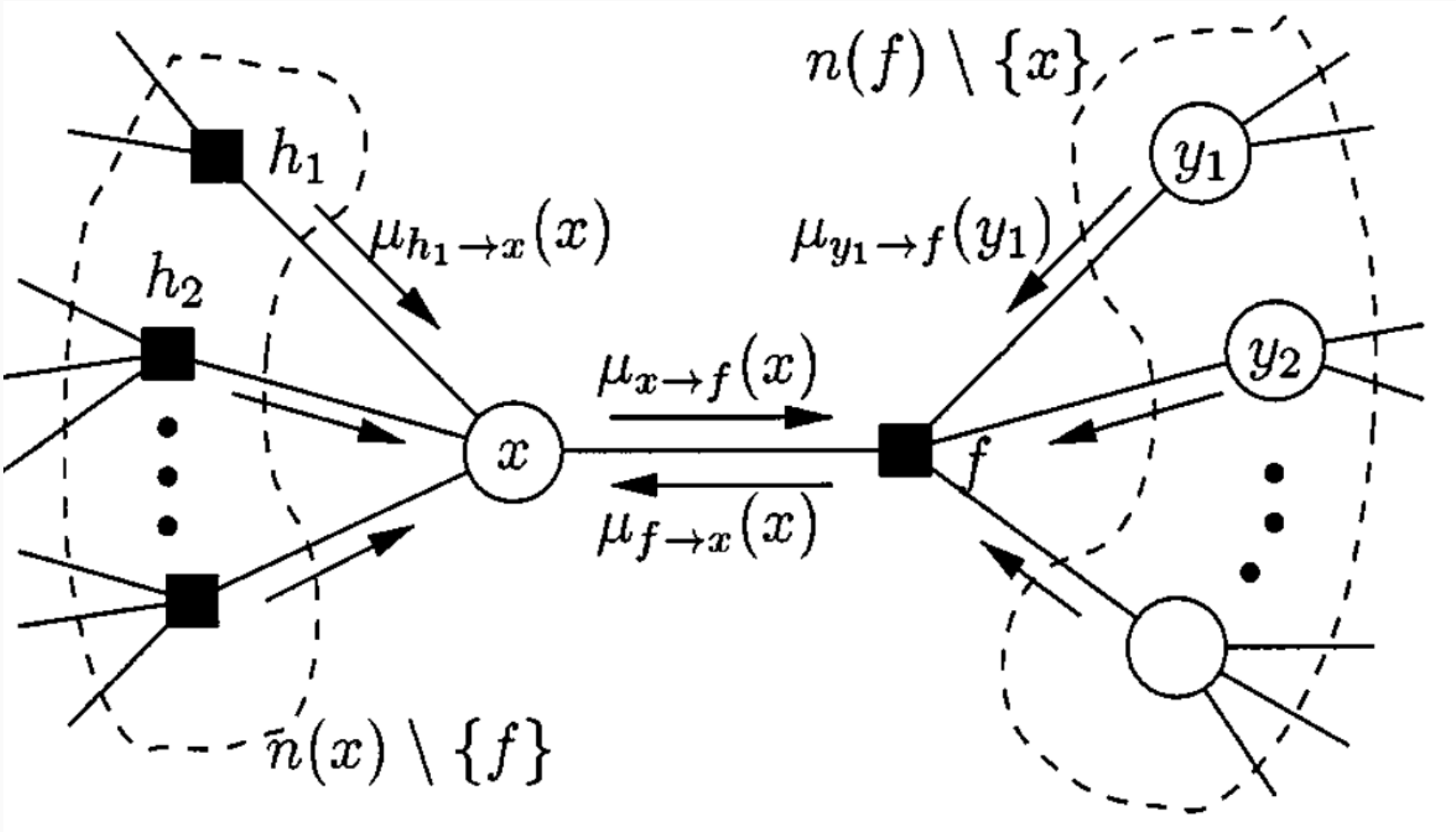
- 15 сентября --- Международный день демократии, который ООН отмечает «с целью повышения информирования общественности», а также Международный день распространения информации о лимфоме
- 15 сентября 1792 г. Наполеон прибыл в Аяччо и возглавил батальон волонтёров, 15 сентября 1805 г. Россия объявила Наполеону первую войну, а 15 сентября 1812 г. Наполеон вошёл в Московский Кремль

SUM--PRODUCT И MAX--SUM

АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

- Чтобы обобщить, удобно рассмотреть опять фактор-граф.
- Предположим, что фактор-граф – дерево (если не дерево, так просто не сработает).
- Алгоритм передачи сообщений решает задачу маргинализации для функции вида $p(x_1, \dots, x_n) = \prod_s f_s(X_s)$, заданной в виде фактор-графа.
- Передаём сообщения по направлению к нужному узлу от переменных к функциям и наоборот.

ПЕРЕДАЧА СООБЩЕНИЙ



АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

- Чтобы найти $p(x_k)$, запишем $p(x_1, \dots, x_n) = \prod_{s \in \neq(x_k)} F_s(x_k, X_s)$, где X_s – переменные из поддерева с корнем в f_s. Тогда

$$p(x_k) = \sum_{x_{i \neq k}} p(x_1, \dots, x_n) = \prod_{s \in \neq(x_k)} \left[\sum_{X_s} F_s(x_k, X_s) \right] =$$
$$= \prod_{s \in \neq(x_k)} \mu_{f_s \to x_k}(x_k),$$

где $\mu_{f_s \to x_k}(x_k)$ – сообщения от соседних функций к переменной x_k.

АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

- Чтобы найти $\mu_{f_s \to x_k}(x_k)$, заметим, что $F_s(x_k, X_s)$ тоже можно разложить по соответствующему подграфу:

$$F_s(x_k, X_s) = f_s(x_k, Y_s) \prod_{y \in Y_s} G_y(y, X_{s,y}),$$

  где Y_s – переменные, непосредственно связанные с f_s (кроме x_k), $X_{s,y}$ – соответствующие поддеревья.
- Итого получаем

$$\mu_{f_s \to x_k}(x_k) = \sum_{Y_s} f_s(x_k, Y_s) \prod_{y \in Y_s} \left(\sum_{X_{s,y}} G_y(y, X_{s,y}) \right) = \\ = \sum_{Y_s} f_s(x_k, Y_s) \prod_{y \in Y_s} \mu_{y \to f_s}(y).$$

АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

- Можно аналогично подсчитать, что $\mu_{y \to f_s}(y) = \prod_{f \in \neq(y)\ f_s} \mu_{f \to y}(y)$.

АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

- Итак, получился простой и понятный алгоритм:
  - как только узел получил сообщения от всех соседей, кроме одного, он сам начинает передавать сообщение в этого соседа;
  - сообщение по ребру между функцией и переменной является функцией от этой переменной;
  - узел-переменная x передаёт сообщение

$$\mu_{x\to f}(x) = \prod_{g\in\neq(x)\ f} \mu_{g\to x}(x);$$

  - узел-функция $f(x, Y)$ передаёт сообщение

$$\mu_{f\to x}(x) = \sum_{y\in Y} f(x, Y) \prod_{y\in Y} \mu_{y\to f}(y);$$

  - начальные сообщения в листьях $\mu_{x\to f}(x) = 1$, $\mu_{f\to x}(x) = f(x)$.

АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

- Когда сообщения придут из всех соседей в какую-то переменную x_k, можно будет подсчитать

$$p(x_k) = \prod_{f \in \neq(x_k)} \mu_{f \to x_k}(x_k).$$

- Когда сообщения придут из всех соседей в какой-то фактор $f_s(X_s)$, можно будет подсчитать совместное распределение

$$p(X_s) = f_s(X_s) \prod_{y \in \neq(f_s)} \mu_{y \to f_s}(y).$$

- За два прохода (по каждому ребру туда и обратно) можно будет подсчитать маргиналы во всех узлах.

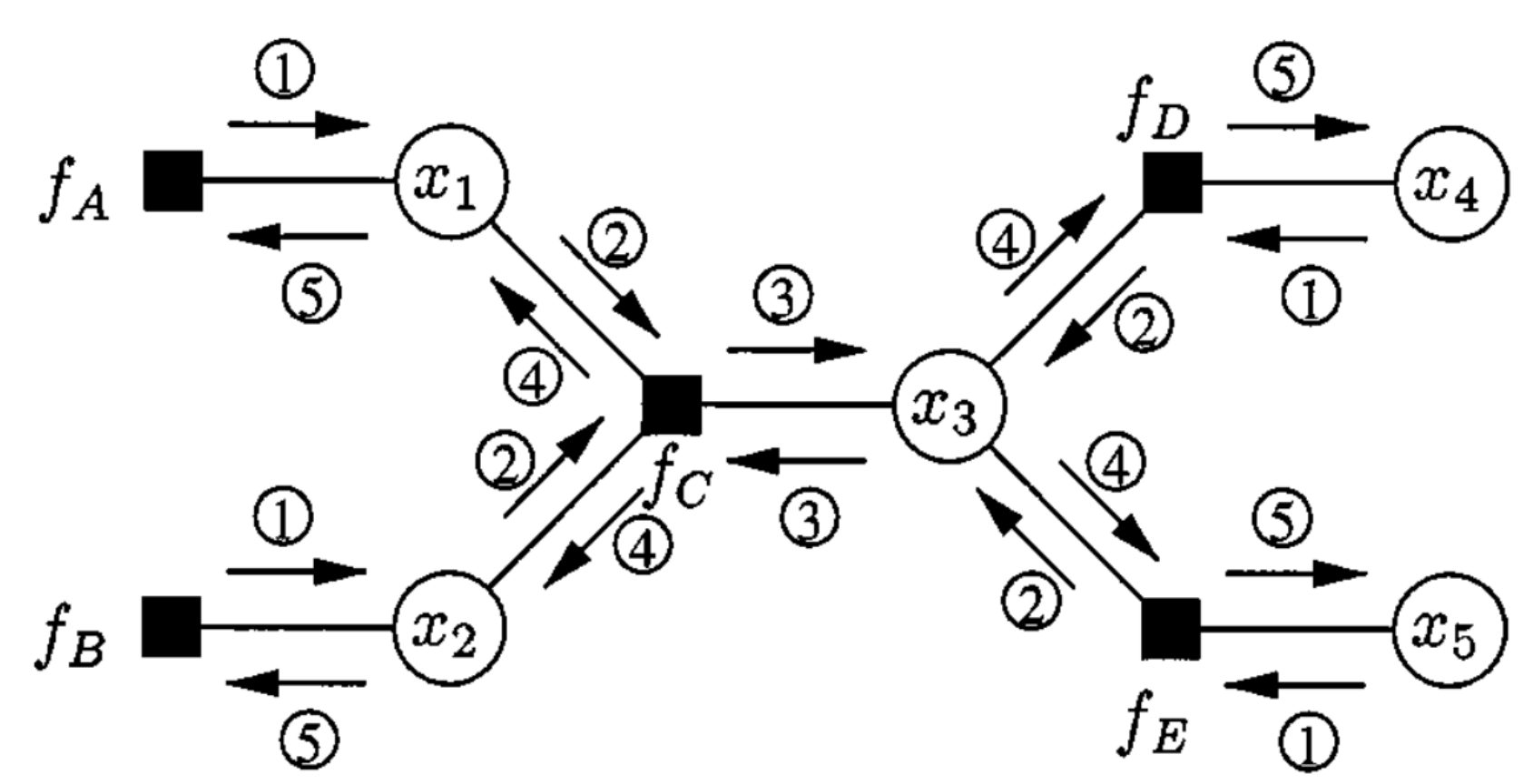
АЛГОРИТМ ПЕРЕДАЧИ СООБЩЕНИЙ

- Это называется алгоритм sum-product, потому что сообщение вычисляется как

$$\mu_{f\to x}(x) = \sum_{y\in Y} f(x, Y) \prod_{y\in Y} \mu_{y\to f}(y).$$

- Задача максимизации $\arg\max_x p(x_1, \dots, x_n)$ решается так же, но алгоритмом max-sum: сумма заменяется на максимум, а произведение на сумму.

ПЕРЕДАЧА СООБЩЕНИЙ



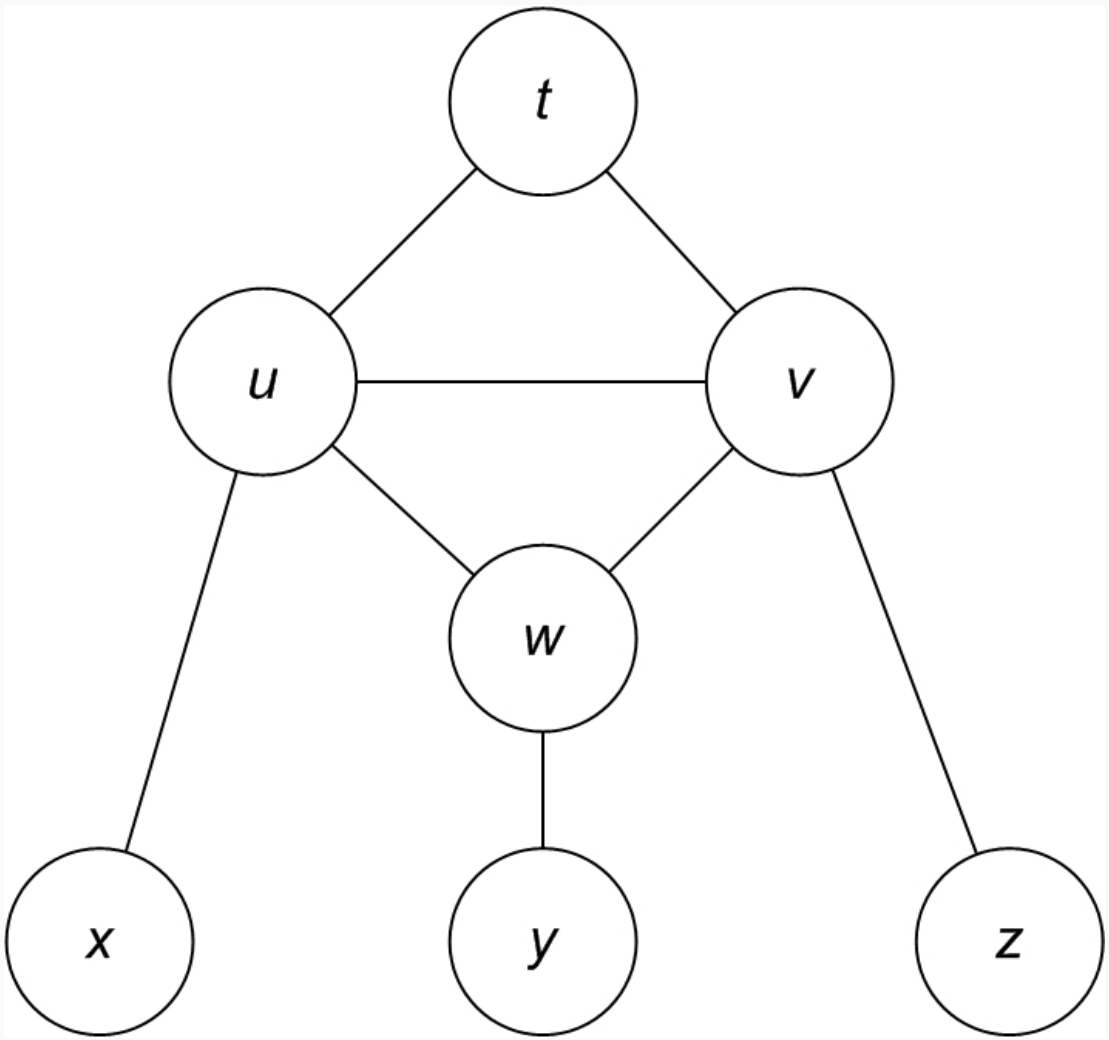
Так что же делать с байесовской сетью?

Для модели не в виде фактор-графа надо просто представить её в виде фактор-графа тем или иным способом.

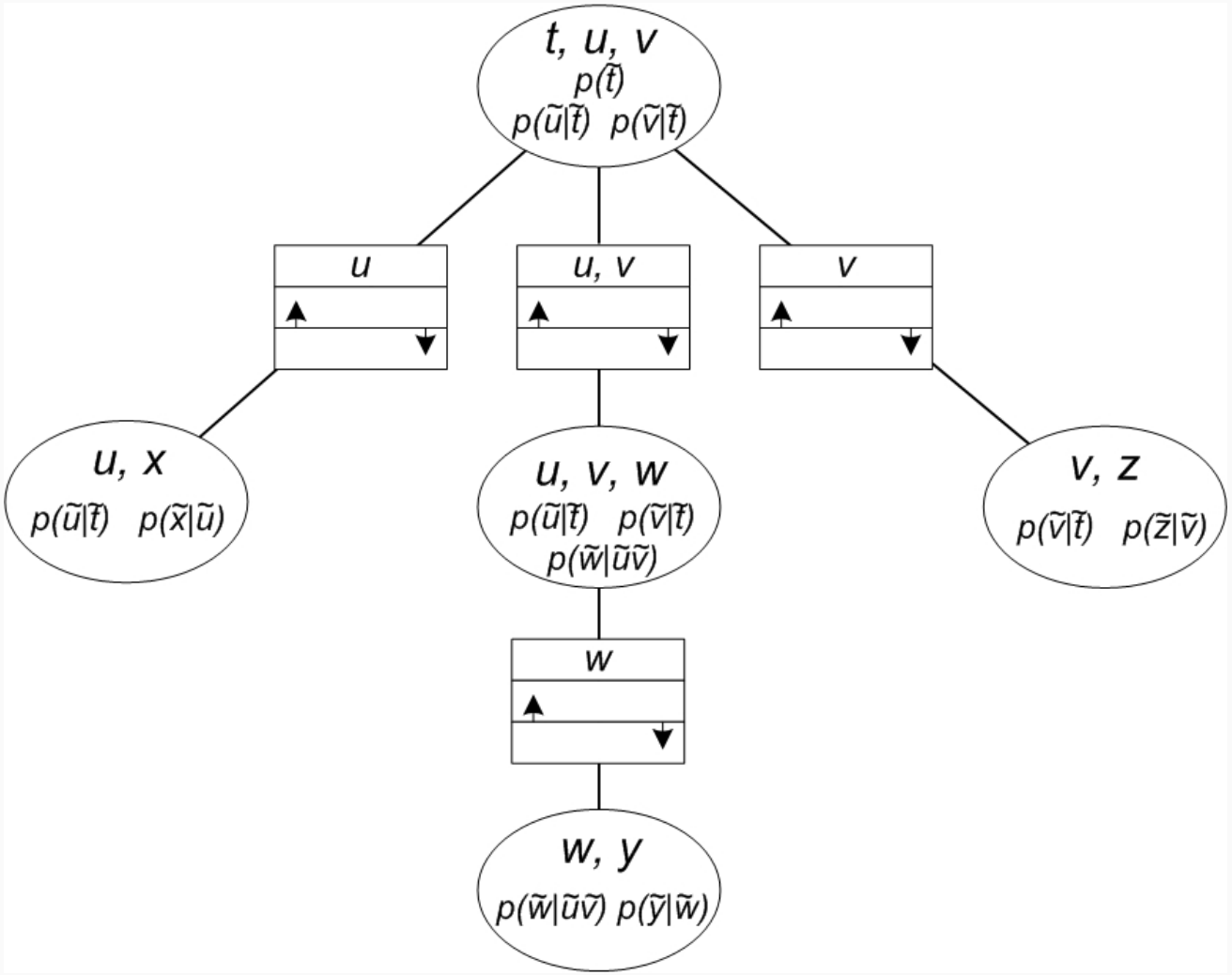
Для байесовской сети это значит, что её сначала нужно представить в виде марковской сети, а потом добавить факторы в явном виде.

ТАК ЧТО ЖЕ ДЕЛАТЬ С БАЙЕСОВСКОЙ СЕТЬЮ?

ТАК ЧТО ЖЕ ДЕЛАТЬ С БАЙЕСОВСКОЙ СЕТЬЮ?



ТАК ЧТО ЖЕ ДЕЛАТЬ С БАЙЕСОВСКОЙ СЕТЬЮ?



THANK YOU!

Thank you for your attention!