# Лекция 2. Графические модели. Общее представление

А. С. Конушин[1] Д. П. Ветров[2] Д. А. Кропотов[3]
В. С. Конушин[1] О. В. Баринова[1]

[1]МГУ, ВМиК, лаб. КГ [2]МГУ, ВМиК, каф. ММП

[2]ВЦ РАН

Спецкурс «Структурные методы анализа изображений и сигналов»

# План

Ликбез
- Условная независимость случайных величин

Графические модели
- Задачи со структурными ограничениями
- Основные проблемы в анализе графических моделей

Байесовские сети
- Факторизация байесовских сетей
- Три элементарных графа
- Пример использования

Марковские сети
- Потенциалы и энергия клик
- Пример использования
- Связь с байесовскими сетями

# План



Ликбез
- Условная независимость случайных величин

Графические модели
- Задачи со структурными ограничениями
- Основные проблемы в анализе графических моделей

Байесовские сети
- Факторизация байесовских сетей
- Три элементарных графа
- Пример использования

Марковские сети
- Потенциалы и энергия клик
- Пример использования
- Связь с байесовскими сетями

# Условная независимость случайных величин

- Случайные величины $x$ и $y$ называются условно независимыми от $z$, если

$$p(x,y|z) = p(x|z)p(y|z)$$

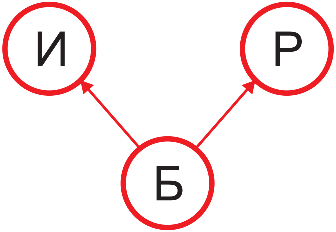
- Другими словами вся информация о взаимозависимостях между $x$ и $y$ содержится в $z$
- Заметим, что из безусловной независимости не следует условная и наоборот
- Основное свойство условно независимых случайных величин

$$p(z|x,y) = \frac{p(x,y|z)p(z)}{p(x,y)} = \frac{p(x|z)p(y|z)p(z)}{p(x,y)} =$$

$$\frac{p(x|z)p(z)p(y|z)p(z)}{p(x,y)p(z)} = \frac{p(z|x)p(z|y)}{p(z)p(x)p(y)p(x,y)} = \frac{1}{Z}\frac{p(z|x)p(z|y)}{p(z)}$$

# Пример

- Рассмотрим следующую гипотетическую ситуацию: римские легионы во главе с императором атакуют вторгшихся варваров
- События «гибель императора» и «уничтожение Рима» не являются независимыми
- Однако, если нам дополнительно известен исход битвы с варварами, эти два события становятся независимыми
- В самом деле, если легионы битву проиграли, то судьба Рима мало зависит от того, был ли император убит в сражении



# План



Ликбез
- Условная независимость случайных величин

Графические модели
- Задачи со структурными ограничениями
- Основные проблемы в анализе графических моделей

Байесовские сети
- Факторизация байесовских сетей
- Три элементарных графа
- Пример использования

Марковские сети
- Потенциалы и энергия клик
- Пример использования
- Связь с байесовскими сетями

# Классическая задача машинного обучения

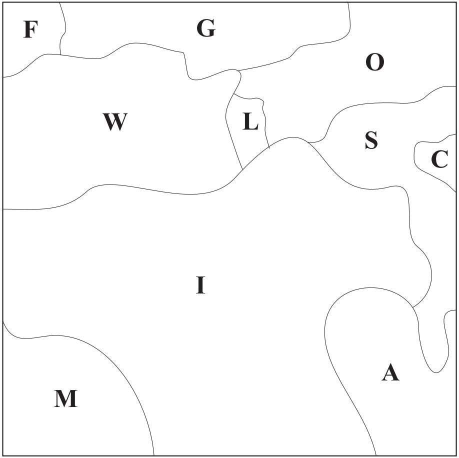
- Задачу машинного обучения можно трактовать как восстановление неизвестных зависимостей между наблюдаемымми переменными $X$ и скрытыми (латентными) переменными $T$. В случае обучения с учителем такое восстановление производится по обучающей выборке $Y$
- В классических задачах машинного обучения предполагается, что обучающая выборка сформирована из однородных и независимых объектов $Y = \{(\boldsymbol{x}_i, \boldsymbol{t}_i)\}_{i=1}^n$
- До недавнего времени вероятностные методы обработки данных ограничивались только таким простейшим случаем, а изложение каждого метода начиналось со слов «Предположим, что нам дана выборка из независимых одинаково распределенных случайных величин...»

# Задачи со структурными ограничениями

- Во многих задачах взаимосвязи между наблюдаемыми и скрытыми переменными носят сложный характер
- В частности, между отдельными переменными существуют вероятностные зависимости
- Факт зависимости переменных друг от друга удобно отображать с помощью неориентированного графа (марковской сети)
- Если связи между переменными причинно-следственные, то их удобно отображать в виде ориентированных графов (байесовских сетей)
- Основным средством работы с графическими моделями служит аппарат теории вероятностей, в ее байесовской интерпретации

# Простой пример

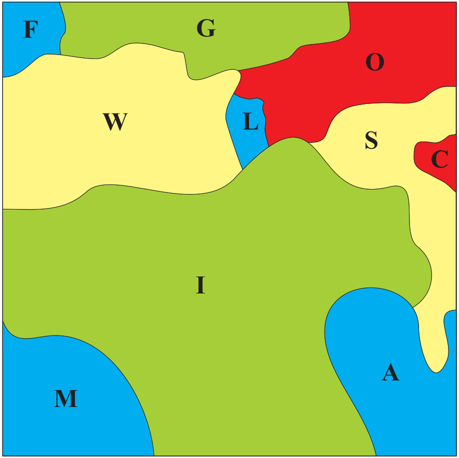
Задача о раскраске областей на плоскости так, чтобы никакие соседние не были окрашены в одинаковый цвет



# Простой пример

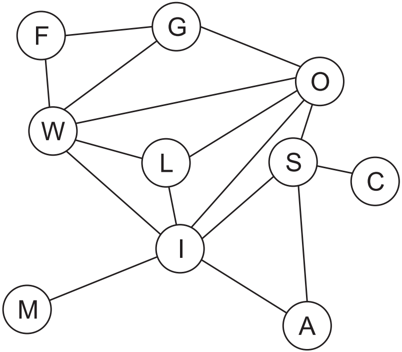


Задача о раскраске областей на плоскости так, чтобы никакие соседние не были окрашены в одинаковый цвет



# Простой пример

Такая задача легко формулируется в терминах графической модели, в которой каждая вершина графа может находиться в одном из четырех состояний




Вопрос залу: почему четырех?

# Примеры задач со структурными связями



- Обработка изображений, сигналов
- Анализ социальных сетей
- Поиск залежей полезных ископаемых
- Анализ естественных языков
- Биомедицина и биоинформатика
- Веб-поиск
- и др.

# План



Ликбез

Условная независимость случайных величин

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

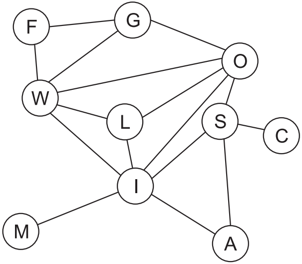
Потенциалы и энергия клик

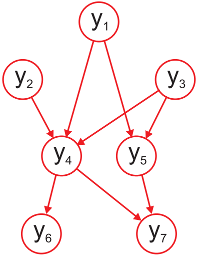
Пример использования

Связь с байесовскими сетями

# Графические модели

- Графическая модель представляет собой ориентированный или неориентированный граф
- Вершины графа соответствуют переменным
- Ребра графа соответствуют вероятностным отношениям, определяющим непосредственные зависимости



# Главные задачи в анализе графических моделей



Обозначим совокупность наблюдаемых переменных $X$, а ненаблюдаемых переменных $T$. Основными задачами в анализе графических моделей являются

- Подсчет условного распределения на значения отдельной скрытой переменной $p(t_i|X)$–?
- Нахождение наиболее вероятной конфигурации скрытых переменных $p(T|X) \rightarrow \max_T$
- Оценка адекватности выбранной графической модели данным $p(X)$–?

# Трудности, возникающие при использовании графических моделей



Ликбез

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Марковские сети

- Не во всех случаях существуют строгие алгоритмы вывода и обучения графических моделей
- Даже там, где они существуют, их применение может оказаться невозможно из-за высоких вычислительных требований и требований к памяти
- В настоящее время в мире активно разрабатываются приближенные эффективные методы обучения и принятия решения в графических моделях
    - Monte Carlo Markov chains (лекция 9)
    - Variational bounds (лекция 5)
    - Expectation propagation (лекция 5)
    - Loopy belief propagation (лекция 4)
    - Tree reweighted message passing (лекция 5))
    - и др.

# План



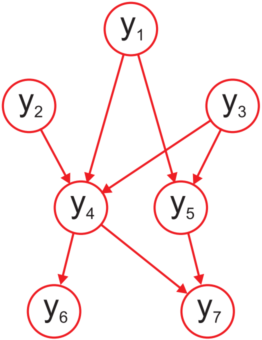
Ликбез
- Условная независимость случайных величин

Графические модели
- Задачи со структурными ограничениями
- Основные проблемы в анализе графических моделей

Байесовские сети
- Факторизация байесовских сетей
- Три элементарных графа
- Пример использования

Марковские сети
- Потенциалы и энергия клик
- Пример использования
- Связь с байесовскими сетями

# Совместное распределение переменных




Совместное распределение системы переменных задается выражением

$$p(Y) = p(y_1, y_2, y_3, y_4, y_5, y_6, y_7) =$$
$$p(y_1)p(y_2)p(y_3)p(y_4|y_1, y_2, y_3)p(y_5|y_1, y_3)p(y_6|y_4)p(y_7|y_4, y_5).$$

# Совместное и условные распределения



- В общем случае совместное распределение для ориентированного графа с $n$ вершинами

$$p(Y) = \prod_{i=1}^{n} p(y_i|\mathrm{pa}_i),$$

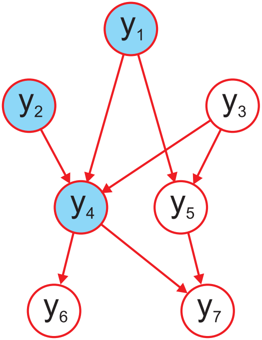
где $\mathrm{pa}_i$ — множество вершин-родителей $y_i$

- Обычно предполагается, что атомарные условные распределения $p(y_i|\mathrm{pa}_i)$ известны
- Зная атомарные распределения, мы можем рассчитать (хотя бы теоретически) любые условные вероятности одних подмножеств переменных по другим подмножествам переменных

# Вычисление условных распределений I



- Вернемся к иллюстрации графической модели из семи переменных
- Пусть нам необходимо найти распределение $(y_5, y_7)$ при заданных значениях $y_1, y_2, y_4$ и неизвестных $y_3, y_6$



# Вычисление условных распределений II

- По определению условной вероятности

$$p(y_5, y_7|y_1, y_2, y_4) = \frac{p(y_1, y_2, y_4, y_5, y_7)}{p(y_1, y_2, y_4)}$$

- Расписываем знаменатель

$$p(y_1, y_2, y_4) = p(y_1)p(y_2)p(y_4|y_1, y_2) = \{Sum\ rule\}$$
$$p(y_1)p(y_2) \int p(y_4|y_1, y_2, y_3)p(y_3)dy_3$$

- Аналогично числитель

$$p(y_1, y_2, y_4, y_5, y_7) = p(y_1)p(y_2)p(y_4|y_1, y_2)p(y_5|y_1)p(y_7|y_5, y_4) = p(y_1)\times$$
$$p(y_2)\left(\int p(y_4|y_1, y_2, y_3)p(y_3)dy_3\right)\left(\int p(y_5|y_1, y_3)p(y_3)dy_3\right)p(y_7|y_5, y_4)$$

- Для взятия возникающих интегралов обычно пользуются различными аппроксимационными методами
- Таким образом, условное распределение выражено через известные атомарные распределения вида $p(y_i|\mathrm{pa}_i)$

# План



Ликбез

Условная независимость случайных величин

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

Потенциалы и энергия клик

Пример использования

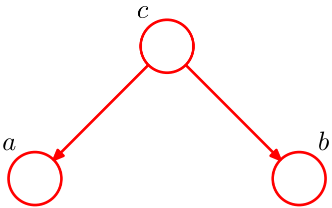
Связь с байесовскими сетями

# Особенности использования байесовских сетей



- По смыслу построения байесовские сети не могут содержать ориентированные циклы, т.к. это будет нарушать правило умножения вероятностей
- Главным достоинством графических моделей является относительно простое выделение условно-независимых величин, которое облегчает дальнейший анализ, позволяя значительно уменьшить количество факторов, влияющих на данную переменную
- В байесовских сетях сделать это несколько сложнее, чем в марковских

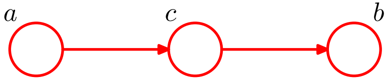
# Граф 1




- Аналогия: Рим (a), император (b) и варвары (c)
- Переменные $a$ и $b$ независимы при заданном $c$
- Возможна маргинализация (исключение переменной)

$$p(a, b) = \int p(a|c)p(b|c)p(c)dc \neq p(a)p(b)$$

# Граф 2




- Аналогия: хорошая работа (a), премия (c), яхта (b)
- Переменные $a$ и $b$ независимы при заданном $c$
- Возможна маргинализация (исключение переменной)

$$p(a,b) = p(a)\int p(b|c)p(c|a)dc \neq p(a)p(b)$$

Ликбез

Графические модели

Байесовские сети

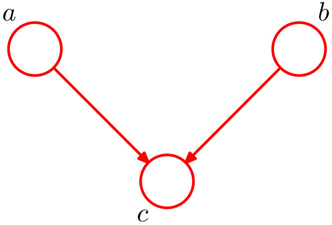
Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

# Граф 3




- Аналогия: вор (a), землетрясение (b) и сигнализация (c)
- Переменные $a$ и $b$ независимы, т.е. $p(a,b)=p(a)p(b)$, но не условно независимы!
- Зависимость $p(c|a,b)$ не может быть выражена через $p(c|a)$ и $p(c|b)$, хотя обратное верно

$$p(c|a)=\int p(c|a,b)p(b)db$$

# План



Ликбез

Условная независимость случайных величин

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

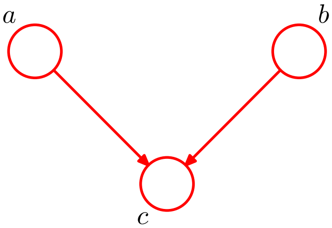
Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

# Пример нестрогих вероятностных рассуждений I






- Рассмотрим последний граф подробнее. Введем обозначения событий: «сигнализация сработала/не сработала» ($s/\neg s$), «вор есть/вора нет» ($v/\neg v$) и «землетрясение произошло/не произошло» ($z/\neg z$)
- Пусть $p(s|v,\neg z) = p(s|v,z) = 1$, $p(s|\neg v, z) = 0.1$, $p(s|\neg v, \neg z) = 0$, $p(v) = 2 \times 10^{-4}$, $p(z) = 10^{-2}$. Графическая модель полностью определена

# Пример нестрогих вероятностных рассуждений II



Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

Допустим, мы получили сигнал тревоги, т.е. имеет место событие $s$. Необходимо оценить вероятность того, что в квартире вор $p(v|s)$

$$p(v|s) = \frac{1}{Z} p(s|v)p(v) = \frac{p(s|v)p(v)}{p(s|v)p(v) + p(s|\neg v)p(\neg v)}$$

$$p(s|\neg v) = p(s|\neg v, z)p(z) + p(s|\neg v, \neg z)p(\neg z) = 10^{-3}$$

$$p(s|v) = 1$$

$$p(v|s) \approx \frac{1}{6}, \quad p(\neg v|s) \approx \frac{5}{6}, \quad Z \approx 1.2 \times 10^{-3}$$

# Пример нестрогих вероятностных рассуждений III



- Пусть теперь дополнительно стало известно, что произошло землетрясение (событие $z$). Как изменится вероятность того, что в квартире вор $p(v|s,z)$?

$$p(v|s,z) = \frac{1}{Z}p(s|v,z)p(v|z), \quad p(v|z) = p(v)$$

$$Z = p(s|v,z)p(v|z) + p(s|\neg v,z)p(\neg v|z) =$$

$$1 \times 2 \times 10^{-4} + 0.1 \times (1 - 2 \times 10^{-4}) = 0.1002$$

$$p(v|s,z) = 0.002, \quad p(\neg v|s,z) = 0.998$$

- Заметим, что события $z$ и $v$ перестали быть независимыми, и добавление сведений о значении $z$ меняет знания о значении $v$. Это называется эффектом оправдания (explaining away)

# Примеры байесовских сетей



Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

- Скрытые марковские модели (лекции 6, 7)
- Фильтр Калмана (лекция 8)
- Экспертные системы
- Вероятностный PCA (лекция 11)
- Смеси экспертов
- Факторный анализ
- и др.

# План



Ликбез
- Условная независимость случайных величин

Графические модели
- Задачи со структурными ограничениями
- Основные проблемы в анализе графических моделей

Байесовские сети
- Факторизация байесовских сетей
- Три элементарных графа
- Пример использования

Марковские сети
- Потенциалы и энергия клик
- Пример использования
- Связь с байесовскими сетями

# Неориентированные графические модели



Ликбез

Графические модели

Байесовские сети

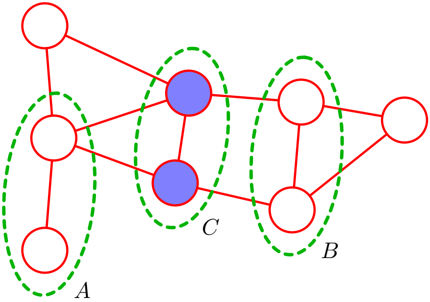
Марковские сети

Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

- При использовании ориентированных графов определение условной независимости не очень просто
- В марковских сетях это проще. На рисунке $A$ и $B$ независимы при условии $C$
- Ребра графа связывают переменные, между которыми существуют непосредственные (а не опосредованные) зависимости



# Факторизация в марковских сетях



Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

- Пусть $y_i$ и $y_j$ независимы при условии, что все остальные переменные нам известны, т.е. $p(y_i, y_j|Y_{\{i,j\}}) = p(y_i|Y_{\{i,j\}})p(y_j|Y_{\{i,j\}})$
- Это означает, что $y_i$ и $y_j$ не соединены ребром (иначе не было бы условной независимости)
- Запишем совместное распределение и применим правило умножения вероятностей

$$p(Y) = p(y_i, y_j|Y_{\{i,j\}})p(Y_{\{i,j\}}) = p(y_i|Y_{\{i,j\}})p(y_j|Y_{\{i,j\}})p(Y_{\{i,j\}})$$

- Таким образом, переменные, не соединенные ребрами, входят в разные множители совместного распределения

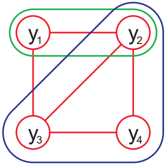
# Потенциалы марковской сети

- В общем виде совместное распределение значений элементов сети записывается с помощью неотрицательных потенциальных функций, определенных на максимальных кликах

$$p(Y) = \frac{1}{Z}\prod_C \psi_C(Y_C), \quad Z = \sum_Y \prod_C \psi_C(Y_C), \quad \psi_C(Y_C) \geq 0$$

- На рисунке синяя клика является максимальной, а зеленая — нет. Совместное распределение имеет вид

$$p(Y) = \frac{1}{Z}\psi_1(y_1, y_2, y_3)\psi_2(y_2, y_3, y_4)$$




Потенциальная функция не обязана иметь вероятностную природу, но чем она больше, тем более вероятны соответствующие значения переменных. Обычно потенциальные функции задаются пользователем исходя из априорных предпочтений тех или иных конфигураций переменных. Реже — настраиваются по данным

# Энергетическая нотация



- Иногда удобно ввести обозначение $\psi_C(Y_C) = \exp(-E_C(Y_C))$, где $E_C(Y_C)$ имеет смысл энергии
- Тогда задача нахождения наиболее вероятного состояния системы сводится к задаче минимизации полной энергии системы

$$\arg\max p(Y) = \arg\max \frac{1}{Z} \prod_C \psi_C(Y_C) =$$

$$\arg\max \exp\left(-\sum_C E_C(Y_C)\right) = \arg\min \sum_C E_C(Y_C)$$

- Заметим, что в отличие от байесовских сетей для полного задания графической модели необходимо знать (или конструктивно уметь подсчитывать) нормировочную константу $Z$

# План



Ликбез
- Условная независимость случайных величин

Графические модели
- Задачи со структурными ограничениями
- Основные проблемы в анализе графических моделей

Байесовские сети
- Факторизация байесовских сетей
- Три элементарных графа
- Пример использования

Марковские сети
- Потенциалы и энергия клик
- Пример использования
- Связь с байесовскими сетями

Ликбез

Графические модели

Байесовские сети

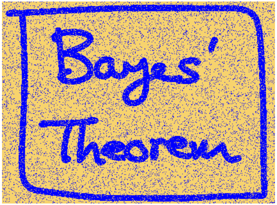
Марковские сети

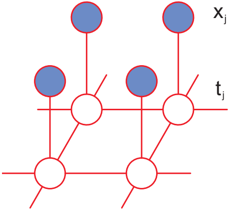
Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

# Фильтрация изображений




- Рассмотрим задачу фильтрации изображения. Пусть $x_i \in \{-1, 1\}$ — наблюдаемые пиксели бинарного изображения, а $t_i \in \{-1, 1\}$ — истинные значения пикселей
- Введем энергию системы

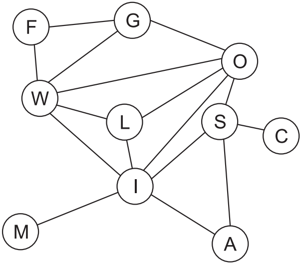
$$E(X, Y) = h \sum_i t_i - \beta \sum_{(i,j) \in E} t_i t_j - \eta \sum_i t_i x_i,$$

где $h \in \mathbb{R}$ позволяет отразить априорные предпочтения в пользу того или иного цвета (например, указать, что желтый цвет встречается чаще, чем синий), $\beta > 0$ выражает степень зависимости между соседними пикселями, а $\eta > 0$ показывает интенсивность шума

# Разметка областей I



Вернемся к примеру со странами

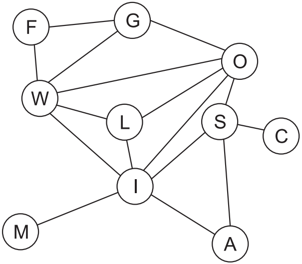



Совместная плотность задается формулой

$$p(X) = \frac{1}{Z}\psi_1(F,G,W)\psi_2(G,O,W)\psi_3(W,O,L,I)\psi_4(I,S,O)\times$$
$$\times\,\psi_5(S,I,A)\psi_6(S,C)\psi_7(M,I)$$

# Разметка областей II

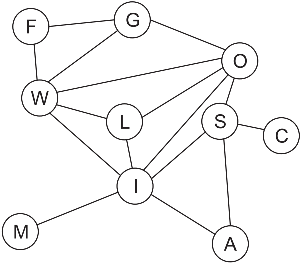





- Предположим, что переменные могут принимать одно из четырех значений $\{red, yellow, blue, green\}$
- Требование несовпадающих цветов областей эквивалентно условию равенства нулю потенциала, если хотя бы два его аргумента имеют одинаковое значение, например $\psi_5(red, blue, red) = 0$

# Разметка областей III

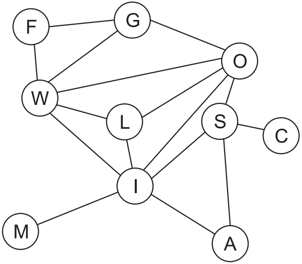





Мы можем снизить число нежелаемых цветовых переходов (например, из желтого в красный) снизив соответствующие значения потенциалов $\psi_7(red, yellow)$, $\psi_7(yellow, red)$, $\psi_6(red, yellow)$ и $\psi_6(yellow, red)$

# Разметка областей IV






Мы можем искусственно способствовать окраске отдельных регионов в выбранные цвета, вводя индивидуальные множители, например

$$\psi_1(F,G,W)=\phi_1(F,G,W)\phi_2(F)\phi_3(W).$$

Теперь можно увеличить значение $\phi_2(blue)$ и $\phi_3(green)$, чтобы получить политическую карту, привычную российскому глазу

# Примеры марковских сетей



Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

- Изображения (лекция 3)
- Социальные сети
- Случайные поля
- Карты сайтов
- и др.

# План



Ликбез
- Условная независимость случайных величин

Графические модели
- Задачи со структурными ограничениями
- Основные проблемы в анализе графических моделей

Байесовские сети
- Факторизация байесовских сетей
- Три элементарных графа
- Пример использования

Марковские сети
- Потенциалы и энергия клик
- Пример использования
- Связь с байесовскими сетями

# Марковские vs. Байесовские сети

Сходства и различия двух типов графических моделей

| Свойство | Марковские сети | Байесовские сети |
|---|---|---|
| Форма | Произв. потенциалов | Произв. потенциалов |
| Потенциалы | Произвольные | Усл. вероятности |
| Циклы | Разрешены | Запрещены |
| Нормировка | $Z =?$ | $Z = 1$ |
| Условная нез-ть | Легко проверяема | Сложнее |
| Полнота | нет | нет |
| Анализ | MCMC, BP, и т.д.. | Сводить к марковским |

# Существующее положение дел

- На сегодняшний день существуют эффективные алгоритмы (sum-product, max-product) анализа ациклических графов (деревьев), решающие все три основные задачи анализа графических моделей
- Частным случаем деревьев являются графы-цепочки, характеризующие, например, сигналы во времени
- В случае наличия циклов сложность точных алгоритмов резко возрастает
- Для анализа графов с циклами в основном используются приближенные методы (loopy BP, EP, MCMC)
- В некоторых частных случаях для анализа циклических сетей существуют эффективные точные алгоритмы, например, разрезы графов (лекция 3)

# Сведение байесовских сетей к марковским



- Наиболее разработаны в настоящее время методы анализа марковских сетей
- Байесовскую сеть можно легко свести к марковской с потерей информации, переженив всех родителей (морализация)
- Заметим, что в приведенном примере все полезные свойства оказались потеряны, и мы получили клику, в которой все зависят от всех