S. Canzar[1] K. Elbassioni[2] A. Elmasry[3] R. Raman[4]

On the Approximability of the Maximum Interval Constrained Coloring Problem

[1]Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

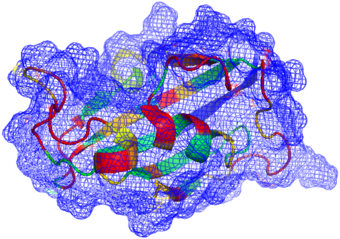
[2]Max-Planck-Institut für Informatik, Saarbrücken, Germany

[3]Datalogisk Institut, University of Copenhagen, Denmark

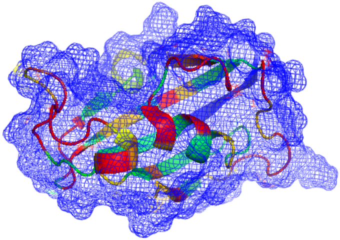
[4]DIMAP and Department of Computer Science, University of Warwick, UK

ISAAC'10

Motivation

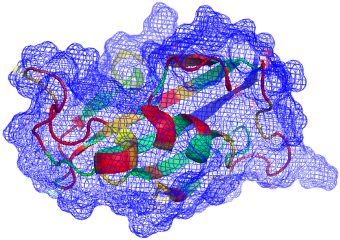


Motivation



- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes. . .)

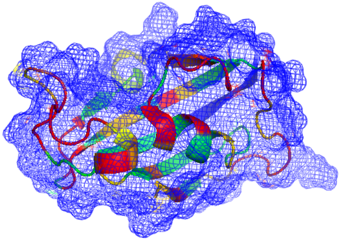
Motivation



- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes...)

⇒ Exchange of labile hydrogens for deuteriums (HDX)

Motivation

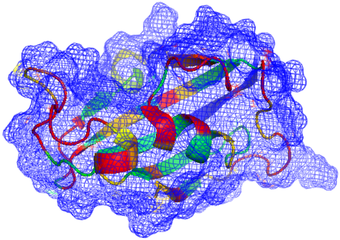


- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes...)

$\Rightarrow$ Exchange of labile hydrogens for deuteriums (HDX)

- exchange rates as measure of solvent accessibility $\rightarrow$ structure!

Motivation

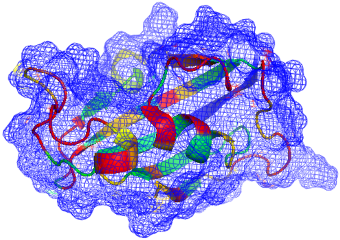


- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes...)

⇒ Exchange of labile hydrogens for deuteriums (HDX)

- exchange rates as measure of solvent accessibility → structure!
- monitored by mass spectrometry (MS) (NMR limited)

Motivation

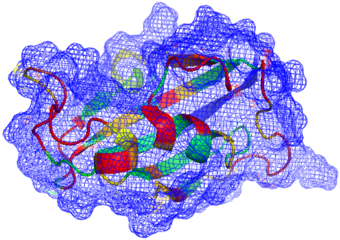


- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes...)

$\Rightarrow$ Exchange of labile hydrogens for deuteriums (HDX)

- exchange rates as measure of solvent accessibility $\rightarrow$ structure!
- monitored by mass spectrometry (MS) (NMR limited)
- output: aggregate exchange data for peptic fragments

Motivation



- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes...)

$\Rightarrow$ Exchange of labile hydrogens for deuteriums (HDX)

- exchange rates as measure of solvent accessibility $\rightarrow$ structure!
- monitored by mass spectrometry (MS) (NMR limited)
- output: aggregate exchange data for peptic fragments

$\Rightarrow$ Increase resolution from fragments to single residues!

Problem Illustration

G	R	Y	R	G	Y	R	R	Y
1	2	3	4	5	6	7	8	9

Problem Illustration

G	R	Y	R	G	Y	R	R	Y
1	2	3	4	5	6	7	8	9

Experiments

$[1, 5]$ — — $(2, 2, 1)$

$[3, 7]$ — — $(2, 1, 2)$

$[6, 8]$ — — $(2, 0, 1)$

$[5, 9]$ — — $(2, 1, 2)$

Problem Illustration

G	R	Y	R	G	Y	R	R	Y
1	2	3	4	5	6	7	8	9

Experiments ↓

$[1,5]$ — — $(2,2,1)$

$[3,7]$ — — $(2,1,2)$

$[6,8]$ — — $(2,0,1)$

$[5,9]$ — — $(2,1,2)$

↑ feasible assignment

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V
- H/D exchange rates $\longmapsto$ Set of color classes $[k]$

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V
- H/D exchange rates $\longmapsto$ Set of color classes $[k]$
- Requirement function $r : \mathcal{I} \times [k] \mapsto \mathbb{N}$

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V
- H/D exchange rates $\longmapsto$ Set of color classes $[k]$
- Requirement function $r : \mathcal{I} \times [k] \mapsto \mathbb{N}$

Feasibility

$\mathcal{I}$ is *feasible* or *colorable* if there exists a coloring $\chi : V \mapsto [k]$ such that for every $I \in \mathcal{I}$ we have $|\{j \in I \mid \chi(j) = c\}| = r(I, c)$.

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V
- H/D exchange rates $\longmapsto$ Set of color classes $[k]$
- Requirement function $r : \mathcal{I} \times [k] \mapsto \mathbb{N}$

Feasibility

$\mathcal{I}$ is *feasible* or *colorable* if there exists a coloring $\chi : V \mapsto [k]$ such that for every $I \in \mathcal{I}$ we have $|\{j \in I \mid \chi(j) = c\}| = r(I, c)$.

MaxFeasibleColoring (MFC)

Given non-negative weights $w : \mathcal{I} \mapsto \mathbb{R}_+$, find a maximum weight colorable subset $\mathcal{I}' \subseteq \mathcal{I}$.

Previous Results & Contribution

- [Althaus et al. SAC'08]: ILP formulation, polynomial for $k = 2$
- [Althaus et al. SWAT'08]:
  - $\mathcal{NP}$-hard in general
  - ± 1 violation by LP rounding
  - $(1 + \epsilon)$ violation in quasi-polynomial time
- [Byrka et al. LATIN'10]: $\mathcal{NP}$-hard/$\mathcal{APX}$-hard for $k = 3$
- [Komusiewicz et al. CPM'09] Fixed-parameter tractable w.r.t. max. interval length, max. no. of intervals containing a vertex.

Previous Results & Contribution

- [Althaus et al. SAC'08]: ILP formulation, polynomial for $k = 2$
- [Althaus et al. SWAT'08]:
  - $\mathcal{NP}$-hard in general
  - ± 1 violation by LP rounding
  - $(1 + \epsilon)$ violation in quasi-polynomial time
- [Byrka et al. LATIN'10]: $\mathcal{NP}$-hard/$\mathcal{APX}$-hard for $k = 3$
- [Komusiewicz et al. CPM'09] Fixed-parameter tractable w.r.t. max. interval length, max. no. of intervals containing a vertex.

Our Results

- $\mathcal{O}(\sqrt{|OPT|})$-approximation algorithm (for constant k)
- MFC is $\mathcal{APX}$-hard for $k = 2$ (reduction from MAX2SAT)

Approximation Algorithm

Idea

Trim solution space and show it still contains a “good” solution.

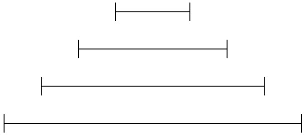
Approximation Algorithm

Idea

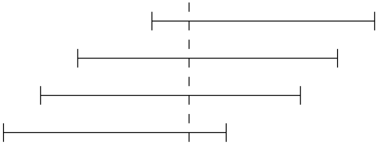
Trim solution space and show it still contains a "good" solution.

(i) Solve easier problem variants optimally:

a) *Tower*



b) *Staircase*



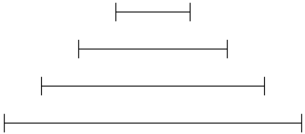
Approximation Algorithm

Idea

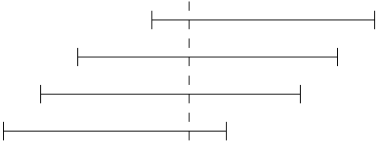
Trim solution space and show it still contains a “good” solution.

(i) Solve easier problem variants optimally:

a) *Tower*



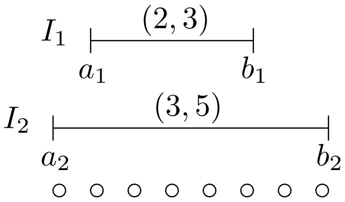
b) *Staircase*



(ii) Pick best (combined) solution

Optimal Tower

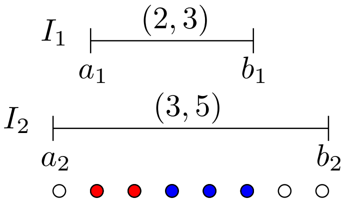
Consider partially ordered set $(\mathcal{P}, \preceq)$:



$$I_1 \preceq I_2 \iff (-a_1, 2, 3, b_1) \leq (-a_2, 3, 5, b_2)$$

Optimal Tower

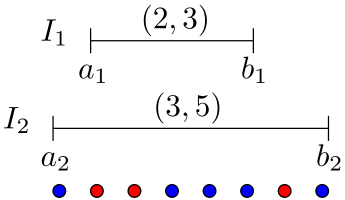
Consider partially ordered set $(\mathcal{P}, \preceq)$:



$$I_1 \preceq I_2 \iff$$
$$(-a_1, 2, 3, b_1) \leq (-a_2, 3, 5, b_2)$$

Optimal Tower

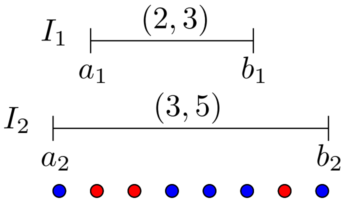
Consider partially ordered set $(\mathcal{P}, \preceq)$:



$$I_1 \preceq I_2 \iff$$
$$(-a_1, 2, 3, b_1) \leq (-a_2, 3, 5, b_2)$$

Optimal Tower

Consider partially ordered set $(\mathcal{P}, \preceq)$:

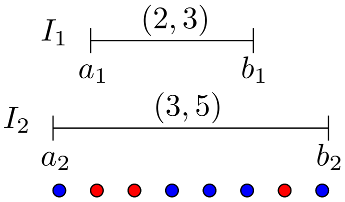


$$I_1 \preceq I_2 \iff (-a_1, 2, 3, b_1) \leq (-a_2, 3, 5, b_2)$$

Observation: $\mathcal{I}' \subseteq \mathcal{I}$ is a feasible tower iff it is a chain in $\mathcal{P}$

Optimal Tower

Consider partially ordered set $(\mathcal{P}, \preceq)$:

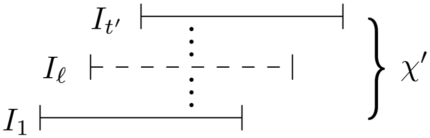


$$I_1 \preceq I_2 \iff$$
$$(-a_1, 2, 3, b_1) \leq (-a_2, 3, 5, b_2)$$

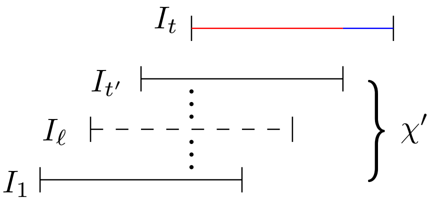
Observation: $\mathcal{I}' \subseteq \mathcal{I}$ is a feasible tower iff it is a chain in $\mathcal{P}$

$\Rightarrow$ Find a maximum-weight chain in $\mathcal{P}$ (in polynomial time)!

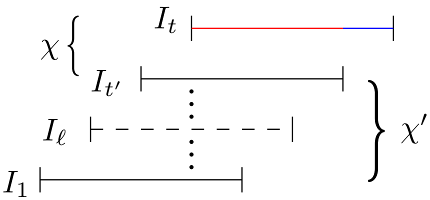
Optimal Staircase



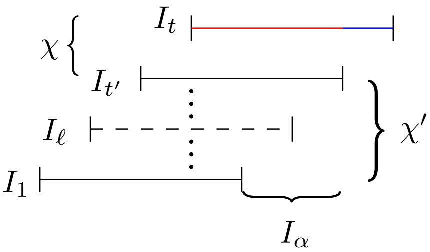
Optimal Staircase



Optimal Staircase

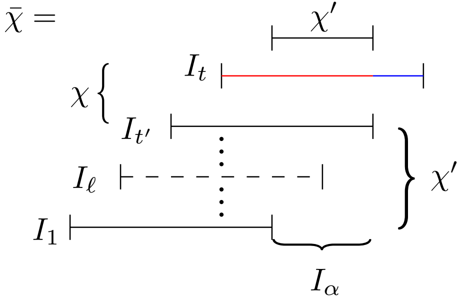


Optimal Staircase



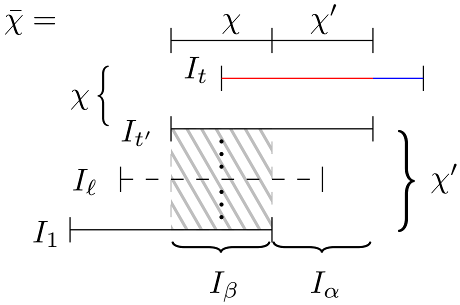
$$|\{v \in I_\alpha \mid \chi(v) = c\}| = |\{v \in I_\alpha \mid \chi'(v) = c\}|,\text{ for all colors } c.$$

Optimal Staircase



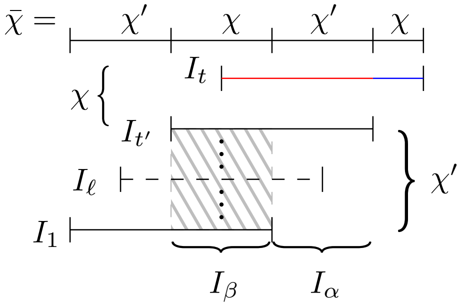
$|\{v \in I_\alpha \mid \chi(v) = c\}| = |\{v \in I_\alpha \mid \chi'(v) = c\}|$, for all colors c.

Optimal Staircase



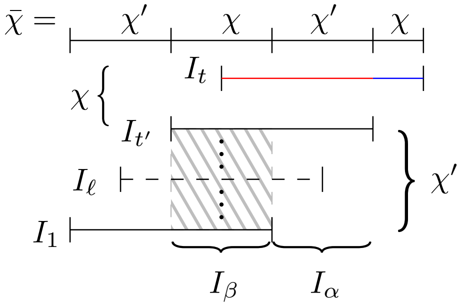
$|\{v \in I_\alpha \mid \chi(v) = c\}| = |\{v \in I_\alpha \mid \chi'(v) = c\}|$, for all colors c.

Optimal Staircase



$|\{v \in I_\alpha \mid \chi(v) = c\}| = |\{v \in I_\alpha \mid \chi'(v) = c\}|$, for all colors c.

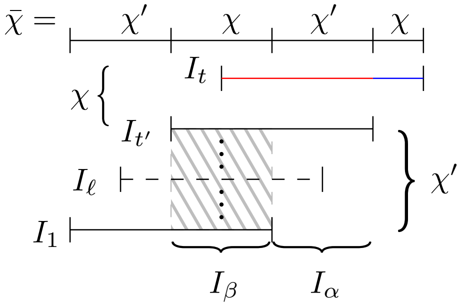
Optimal Staircase



$|\{v \in I_\alpha \mid \chi(v) = c\}| = |\{v \in I_\alpha \mid \chi'(v) = c\}|$, for all colors c.

$D[t', \bar{\mathbf{r}}(I_\alpha)] :=$ weight of optimal solution from I_1 to $I_{t'}$ that can be satisfied by coloring satisfying $\bar{\mathbf{r}}$ on I_α

Optimal Staircase



$$|\{v \in I_\alpha \mid \chi(v) = c\}| = |\{v \in I_\alpha \mid \chi'(v) = c\}|, \text{ for all colors } c.$$

$D[t', \bar{\mathbf{r}}(I_\alpha)] :=$ weight of optimal solution from I_1 to $I_{t'}$ that can be satisfied by coloring satisfying $\bar{\mathbf{r}}$ on I_α

$\Rightarrow$ DP: guess predecessor $I_{t'}$ and "coloring" on I_α

Putting it all together

Define Poset $\mathcal{P} = (OPT, \subseteq)$.

Dilworth Theorem

There is either a chain or an anti-chain in $\mathcal{P}$ of weight at least

$$\frac{w(OPT)}{2\sqrt{|OPT|}}.$$

Putting it all together

Define Poset $\mathcal{P} = (OPT, \subseteq)$.

Dilworth Theorem

There is either a chain or an anti-chain in $\mathcal{P}$ of weight at least

$$\frac{w(OPT)}{2\sqrt{|OPT|}}.$$

- chain (= *tower*):

Putting it all together

Define Poset $\mathcal{P} = (OPT, \subseteq)$.

Dilworth Theorem

There is either a chain or an anti-chain in $\mathcal{P}$ of weight at least

$$\frac{w(OPT)}{2\sqrt{|OPT|}}.$$

- chain (= *tower*): ✓
- anti-chain:

Putting it all together

Define Poset $\mathcal{P} = (OPT, \subseteq)$.

Dilworth Theorem

There is either a chain or an anti-chain in $\mathcal{P}$ of weight at least

$$\frac{w(OPT)}{2\sqrt{|OPT|}}.$$

- chain (= *tower*): ✓
- anti-chain: can be partitioned into 2 sets of independent *staircases* ✓

Putting it all together

Define Poset $\mathcal{P} = (OPT, \subseteq)$.

Dilworth Theorem

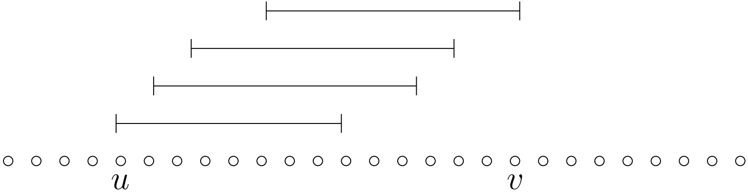
There is either a chain or an anti-chain in $\mathcal{P}$ of weight at least

$$\frac{w(OPT)}{2\sqrt{|OPT|}}.$$

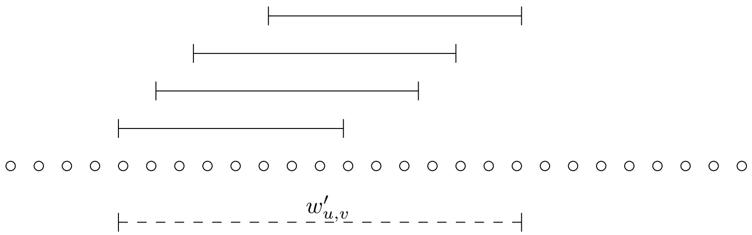
- chain (= *tower*): ✓
- anti-chain: can be partitioned into 2 sets of independent *staircases* ✓

$\Rightarrow$ feasible set of weight at least $\frac{w(OPT)}{4\sqrt{|OPT|}}$

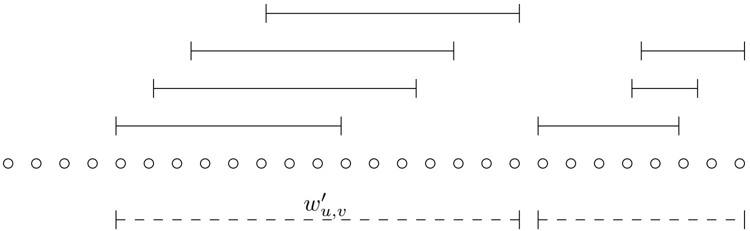
Independent Set of Staircases



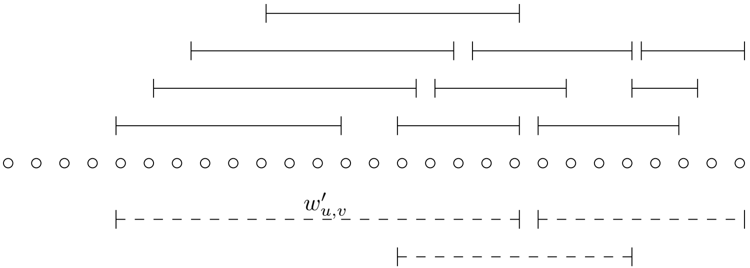
Independent Set of Staircases



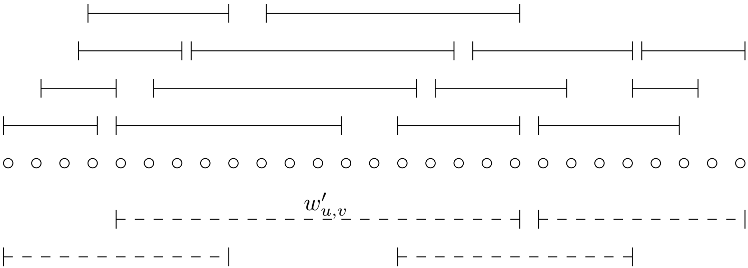
Independent Set of Staircases



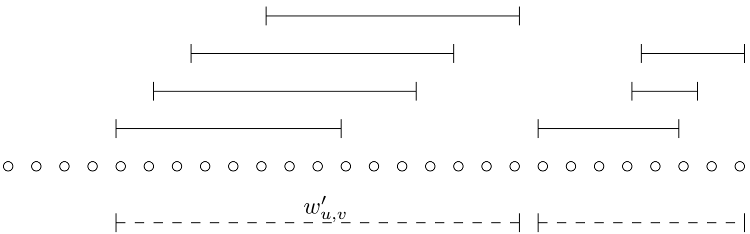
Independent Set of Staircases



Independent Set of Staircases



Independent Set of Staircases



Algorithm

Algorithm 1 $\textsc{MfcApprox}(V, \mathcal{I}, r, w)$

1: $W_1 = \max_{\text{Tower } \mathcal{I}'} w(\mathcal{I}')$
2: **for** every $u, v \in V$ s.t. $u < v$ **do**
3: $\quad w'_{u,v} = \max_{\text{Staircase}(u,v)\ \mathcal{I}'} w(\mathcal{I}')$
4: **end for**
5: Let $W_2 = \max_{\text{Indep.set } \mathcal{I}' \subseteq \{[u,v]:u,v \in V,\ u<v\}} w'(\mathcal{I}')$
6: **return** $\max\{W_1, W_2\}$

Open Questions

- Improve $\mathcal{O}(\sqrt{OPT})$-approximation or hardness?
- Approximation if k part of the input?