

Educating a model student in MOOCs

Tao C. Lee

LCAV, EPFL
TAM, Lausanne, Switzerland
tao.lee@epfl.ch

January 9, 2014

Digital Signal Processing (DSP) MOOC at Coursera

coursera | Global Partners

Courses Partners About ▾ | Tao Lee ▾



Digital Signal Processing

Paolo Prandoni and Martin Vetterli

Learn the fundamentals of digital signal processing theory and discover the myriad ways DSP makes everyday life more productive and fun.



Digital Signal Processing (DSP) MOOC at Coursera

coursera | Global Partners

Courses Partners About ▾ | Tao Lee ▾



Digital Signal Processing

Paolo Prandoni and Martin Vetterli

Learn the fundamentals of digital signal processing theory and discover the myriad ways DSP makes everyday life more productive and fun.



How to analyze and design the DSP MOOC?

- The recap scheduling problem in education

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing
- The design of MDR for the DSP MOOC

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing
- The design of MDR for the DSP MOOC
- Conclusions & future work

Massive Open Online Courses (MOOCs)

- History: 2012, the Year of the MOOC

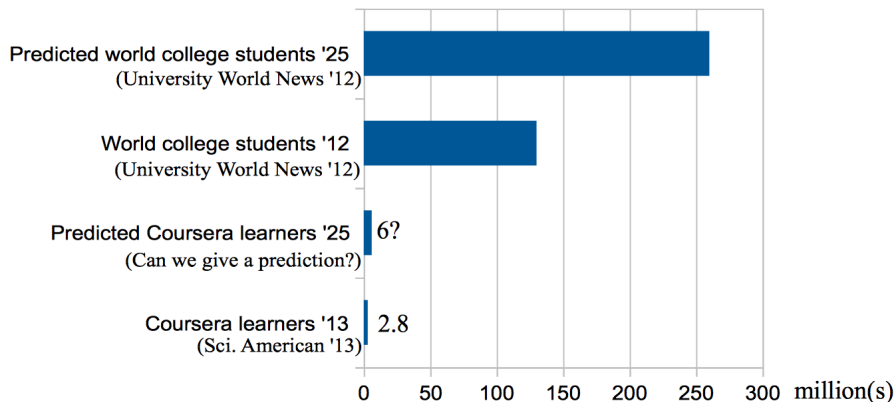
Massive Open Online Courses (MOOCs)

- History: 2012, the Year of the MOOC
- Players: Coursera, edX, Udacity, ...

Massive Open Online Courses (MOOCs)

- History: 2012, the Year of the MOOC
- Players: Coursera, edX, Udacity, ...
- Visibility: > 2.8 million learners on Coursera ('13)

World student population



Research on MOOCs

- Learning sciences community [Blikstein'13] [Ahn'13]

Research on MOOCs

- Learning sciences community [Blikstein'13] [Ahn'13]
- CS community [Novikof'12] [Ghosh'13]

Research on MOOCs

- Learning sciences community [Blikstein'13] [Ahn'13]
- CS community [Novikof'12] [Ghosh'13]
- Others (Science/Nature/NY Times)

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing
- The design of MDR for the DSP MOOC
- Conclusions & future work

The recap scheduling problem in education

Findings of experimental psychology

- Spacing effect: Better to spread studying over time

The recap scheduling problem in education

Findings of experimental psychology

- Spacing effect: Better to spread studying over time
- Need to consider students' needs

The recap scheduling problem in education

Findings of experimental psychology

- Spacing effect: Better to spread studying over time
- Need to consider students' needs
- The timing of review is important!

The recap scheduling problem in education

Findings of experimental psychology

- Spacing effect: Better to spread studying over time
- Need to consider students' needs
- The timing of review is important!

Challenges of MOOCs design

- How to make trade-off between new material and recap?

The recap scheduling problem in education

Findings of experimental psychology

- Spacing effect: Better to spread studying over time
- Need to consider students' needs
- The timing of review is important!

Challenges of MOOCs design

- How to make trade-off between new material and recap?
- How to adapt based on students' feedback?

The recap scheduling problem in education

Findings of experimental psychology

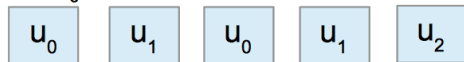
- Spacing effect: Better to spread studying over time
- Need to consider students' needs
- The timing of review is important!

Challenges of MOOCs design

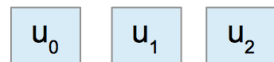
- How to make trade-off between new material and recap?
- How to adapt based on students' feedback?
- Recap schedule needs to satisfy spacing constraints!

Problem illustration

Schedule₀



Schedule₁



Which schedule gives better studying?

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps
- All units have the same (time) length

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps
- All units have the same (time) length
- Students' needs are modeled as: $\{a_k\}$ and $\{b_k\}$

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps
- All units have the same (time) length
- Students' needs are modeled as: $\{a_k\}$ and $\{b_k\}$
- a_k : How long a student should wait before review is beneficial

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps
- All units have the same (time) length
- Students' needs are modeled as: $\{a_k\}$ and $\{b_k\}$
- a_k : How long a student should wait before review is beneficial
- b_k : How long a student can retain learned material

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps
- All units have the same (time) length
- Students' needs are modeled as: $\{a_k\}$ and $\{b_k\}$
- a_k : How long a student should wait before review is beneficial
- b_k : How long a student can retain learned material
- $T_n(k)$: The timestamp of the k^{th} occurrence of unit n

Education of a model student

Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps
- All units have the same (time) length
- Students' needs are modeled as: $\{a_k\}$ and $\{b_k\}$
- a_k : How long a student should wait before review is beneficial
- b_k : How long a student can retain learned material
- $T_n(k)$: The timestamp of the k^{th} occurrence of unit n
- $a_k \leq |T_n(k+1) - T_n(k)| \leq b_k$

Education of a model student

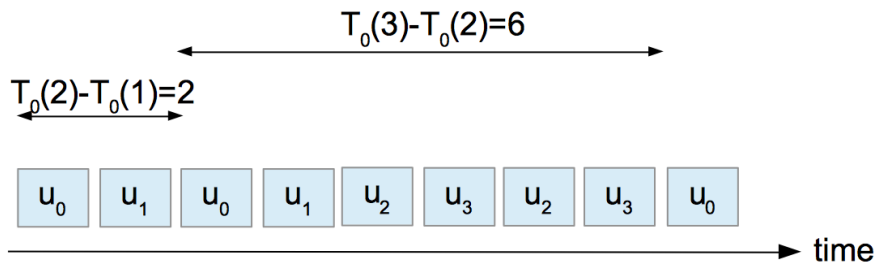
Novikof, Kleinberg, Strogatz '12, PNAS

- Model the education process as a sequence of abstract units
- New material and review are spaced over time
- Students' needs modeled as constraints on the schedules

Models

- $\{u_i\} = u_1, u_2, u_3, u_1$: u_1 intro at 1st and recap at 4th timesteps
- All units have the same (time) length
- Students' needs are modeled as: $\{a_k\}$ and $\{b_k\}$
- a_k : How long a student should wait before review is beneficial
- b_k : How long a student can retain learned material
- $T_n(k)$: The timestamp of the k^{th} occurrence of unit n
- $a_k \leq |T_n(k+1) - T_n(k)| \leq b_k$
- $a_k \leq a_{k+1}, b_k \leq b_{k+1} \forall k$

Model illustration



Infinite Perfect Learning (IPL): $\exists\{u_i\}$ satisfy $(a_k, b_k) \forall i$ as $i \rightarrow \infty$

Infinite Perfect Learning (IPL): $\exists\{u_i\}$ satisfy $(a_k, b_k) \forall i$ as $i \rightarrow \infty$

Introduction time function $T_n(1)$: the first timestep u_n is introduced

Infinite Perfect Learning (IPL): $\exists\{u_i\}$ satisfy $(a_k, b_k) \forall i$ as $i \rightarrow \infty$

Introduction time function $T_n(1)$: the first timestep u_n is introduced

Model	(a_k, b_k)	Asymptotics of $T_n(1)$
Fixed length recap schedule	$(2^k, 2^{k-1}(k+1))$	$\Theta(n \log_2 n)$
Slow flashcard schedule	(k, k^2)	$\Theta(n^2)$
Flexible students	$(1, k+1)$	$\Theta(n^2)$
Finicky slow students	(k, k)	None

The Fixed Length Recap (FLR) schedule

$(a_k, b_k) = (2^k, 2^{k-1}(k+1))$, $k = 2$, first 12 units,

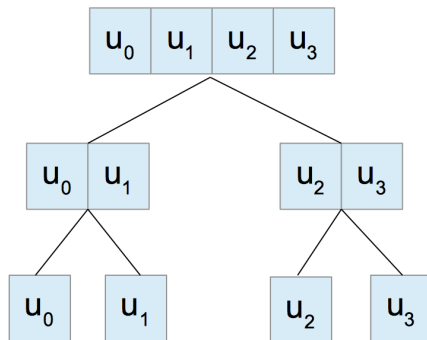
$u_0, u_1, u_0, u_1, u_2, u_3, u_2, u_3, u_0, u_1, u_2, u_3.$

The Fixed Length Recap (FLR) schedule

$(a_k, b_k) = (2^k, 2^{k-1}(k+1))$, $k = 2$, first 12 units,

$u_0, u_1, u_0, u_1, u_2, u_3, u_2, u_3, u_0, u_1, u_2, u_3$.

DFS traversal



Problems

- Recap units are usually shorter than new units
- Not all units need recap
- Nonlinear asymptotics: Bad for course planning

Problems

- Recap units are usually shorter than new units
- Not all units need recap
- Nonlinear asymptotics: Bad for course planning

Can we do better?

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing
- The design of MDR for the DSP MOOC
- Conclusions & future work

The Monotonically Decreasing Recap (MDR) schedule

$(a_k, b_k) = (2^k, 2^{k-1}(k+1))$, $k = 2$, decreasing factor = 0.5, first 16 units,

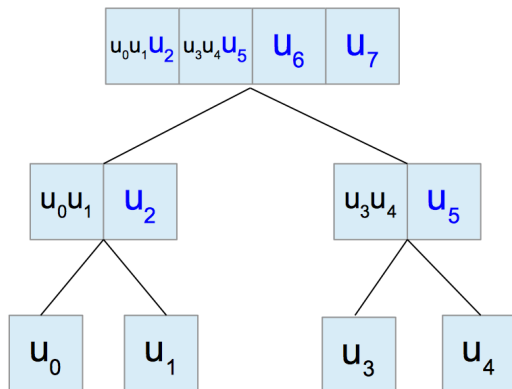
$u_0, u_1, u_0, u_1, u_2, u_3, u_4, u_3, u_4, u_5, u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7$.

The Monotonically Decreasing Recap (MDR) schedule

$(a_k, b_k) = (2^k, 2^{k-1}(k+1))$, $k = 2$, decreasing factor = 0.5, first 16 units,

$u_0, u_1, u_0, u_1, u_2, u_3, u_4, u_3, u_4, u_5, u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7$.

DFS traversal



Theorem

In the MDR schedule with a decreasing factor $\alpha \leq 0.5$,

$T_n(1)$ grows as $\Theta(n)$.

Proof: Details see report.

$$n \leq T_n(1) \leq 2 \cdot n$$

Asymptotics of $T_n(1)$ for $0.5 < \alpha < 1$ is under investigation ...

Theorem

In the MDR schedule with a decreasing factor $\alpha = 0.5$, the units introduced in the leaf nodes adheres to the spacing constraints

$$a_k = 2^k$$

$$b_k = 2^{k-1}(k + 1).$$

under the condition that the units introduced in the nonleaf nodes can be arbitrarily distributed in the nodes they are present.

Proof: See report.

Properties

- Review units are shorter
- Only leaf units satisfy spacing constraints
- Linear schedule

Properties

- Review units are shorter
- Only leaf units satisfy spacing constraints
- Linear schedule

How to design a MDR for the DSP MOOC?

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing
- The design of MDR for the DSP MOOC
- Conclusions & future work

Data at hand

- Online forums
- Homework, labs, quizzes, exams
- Survey

Data at hand

- Online forums
- Homework, labs, quizzes, exams
- Survey

Online forums

- Direct feedback from students
- Collaboration between students and teaching staffs

$$\begin{aligned} \text{forum data} &= \{ \text{thread}_i \} \\ \text{thread}_i &= \{ \text{post}_j \} \\ \text{post}_j &= \text{post}_j | \{ \text{comment}_k \}. \end{aligned}$$

Concepts of interest in the forum

- Supervised learning approach
- Machine learning software: Kea [Witten'98]
- Human screening of extracted keywords

Concept	Keyphrases
Signal	Gaussian, sinc, ...
Basic math	summation, polynomial, ...
Filter	filter, convolution, FIR, ...
Fourier analysis	Fourier, DFS, DFT, ...
Periodicity	periodic, period
Vector space	Hilbert space, basis, subspace, ...
Sampling	sampling frequency/theorem
Orthogonality	orthonormal, orthogonal
Interpolation	Lagrange/linear interpolation, ...
Wavelet	wavelet, wavelets

Another way to look at forum data

Sequence of posts with tags

$forum\ data = \{p_1, p_2, \dots, p_n\}$,
where $p_k = p_k(timestamp, participant, keyphrases)$.

Another way to look at forum data

Sequence of posts with tags

$$\text{forum data} = \{p_1, p_2, \dots, p_n\},$$

where $p_k = p_k(\text{timestamp}, \text{participant}, \text{keyphrases})$.

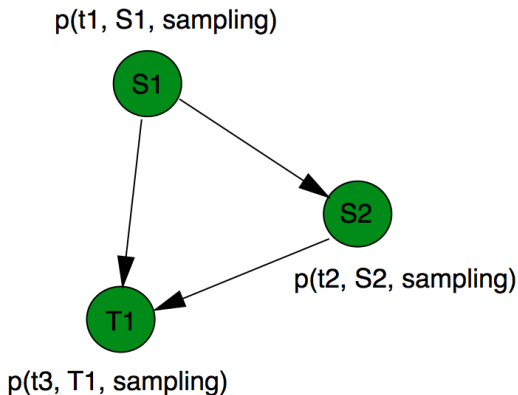
A forum thread,

$$p_1(t_1, S_1, \text{key}), p_2(t_2, S_2, \text{key}), p_3(t_3, T_1, \text{key}),$$

where $t_1 < t_2 < t_3$, S_1, S_2 are two students
 T_1 is a teacher, $\text{key} = \text{sampling}$

Graph of forum data

More visualization (edges follow chronological order) ...



T1 is an authoritative source ...

Filtering the seq of forum posts

Raw seq not informative ...

Definition

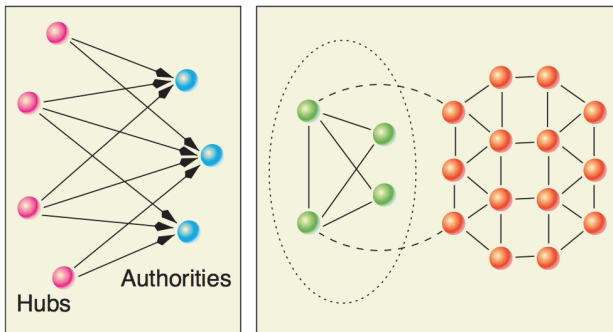
A filtering operation F on the sequence of forum posts $\{p_k\}$, ranks and keeps the n_{top} top-ranked p_k as $\{p'_k\}$

$$F(\{p_k\}, n_{top}) = \{p'_k\}.$$

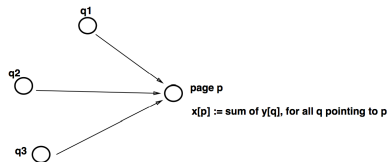
Ex: timestamp, authority, concept filtering ...

HITS algorithm

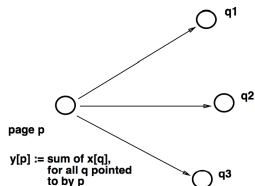
Finding authoritative sources on the web [Kleinberg'01]



HITS algorithm



(c) I Operator



(d) O Operator

HITS algorithm

HITS(G, k)

G : a collection of n linked pages

k : a natural number

Let z denote the vector $(1,1,1,\dots,1) \in \mathbb{R}^n$. Set $x_0 \leftarrow z$.

Set $y_0 \leftarrow z$.

For $i = 1, 2, \dots, k$

 Apply the I operation to (x_{i-1}, y_{i-1}) , obtaining new x -weights x_i' .

 Apply the O operation to (x_i', y_{i-1}) , obtaining new y -weights y_i' .

 Normalize x_i' , obtaining x_i .

 Normalize y_i' , obtaining y_i .

End

Return (x_k, y_k) .

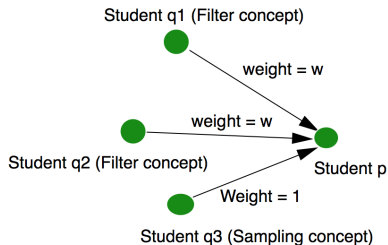
Definition

An authority filtering operation AF on the sequence of forum posts $\{p_k\}$, ranks and keeps the chronologically ordered posts of the n_{top} top-ranked nodes as $\{p'_k\}$ by authority weights using HITS Algorithm,

$$AF(\{p_k\}, n_{top}) = \{p'_k\}.$$

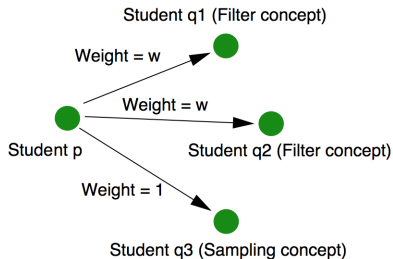
Weighted-HITS algorithm

Emphasize the nodes with posts containing keyphrase K ...



$$x[p] := \text{Sum } y[q] * \text{Weight for all } q \text{ point to } p$$

(g) I' Operator



$$y[p] := \text{Sum } x[q] * \text{Weight for all } q \text{ pointed by } p$$

(h) O' Operator

Weighted-HITS algorithm

Given weights $\{x'_p\}$, $\{y'_p\}$, an amplification factor a ($a \geq 1$), the I' operation updates the x-weights as follows.

$$x'_p \leftarrow \sum_{q:(q,p) \in E} \theta(K) \cdot y'_q.$$

The O' operation updates the y-weights as follows.

$$y'_p \leftarrow \sum_{q:(p,q) \in E} \theta(K) \cdot x'_q,$$

where

$$\theta(K) = \begin{cases} a & : p \text{ or } q \text{ has posts containing keyphrase } K \\ 1 & : \text{otherwise} \end{cases}$$

Weighted-HITS algorithm

WeightedHITS (G, k, K)

G : a collection of n linked nodes

k : a natural number

K : a keyphrase

Let z denote the vector $(1, 1, 1, \dots, 1) \in R_n$. Set $x_0 \leftarrow z$.

Set $y_0 \leftarrow z$.

For $i = 1, 2, \dots, k$

Apply the I' operation to (x_{i-1}, y_{i-1}) , obtaining new x -weights $x'_{i,new}$.

Apply the O' operation to $(x'_{i,new}, y_{i-1})$, obtaining new y -weights $y'_{i,new}$.

Normalize $x'_{i,new}$, obtaining x'_i .

Normalize $y'_{i,new}$, obtaining y'_i .

End

Return (x'_k, y'_k) .

Theorem

The sequences x'_1, x'_2, x'_3, \dots and y'_1, y'_2, y'_3, \dots converge (to limits x'^ and y'^* respectively)*

Proof: See report.

Definition

A concept filtering operation CF on the sequence of forum posts $\{p_k\}$, ranks and keeps the chronologically ordered posts of the n_{top} top-ranked nodes as $\{p'_k\}$ by authority weights using keyphrase K and Weighted-HITS algorithm,

$$CF(\{p_k\}, n_{top}, K) = \{p'_k\}.$$

Theorem

Authority filtering and concept filtering are nonlinear operations.

Proof: See report.

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing
- The design of MDR for the DSP MOOC
- Conclusions & future work

Education units of the DSP MOOC (Spr '13)

17 units in 9 weeks ...

Unit	Description
u_0	Introduction to DSP, basic mathematics
u_1	Hilbert space and approximation
u_2	Introduction to DFT
u_3	DFT examples
u_4	DFT, DFS, DTFT
u_5	Relationship between transforms
u_6	Linear filters
u_7	Frequency response
u_8	Realizable filters
u_9	Filter design

Education units of the DSP MOOC (Spr '13)

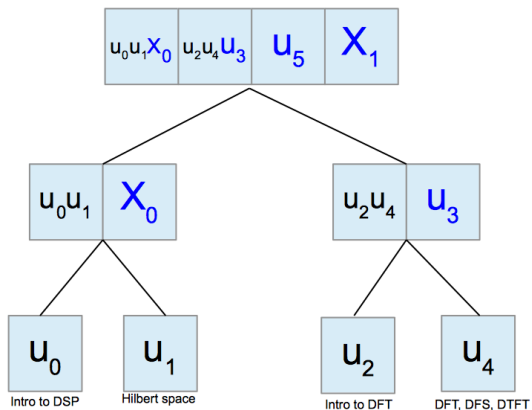
17 units in 9 weeks ...

Unit	Description
u_{10}	Interpolation and bandlimited signals
u_{11}	Sampling and aliasing
u_{12}	Stochastic signal processing
u_{13}	Image processing
u_{14}	Image filtering
u_{15}	Digital communication systems
u_{16}	Modulation and demodulation

Divide and conquer: 3 MDR schedules

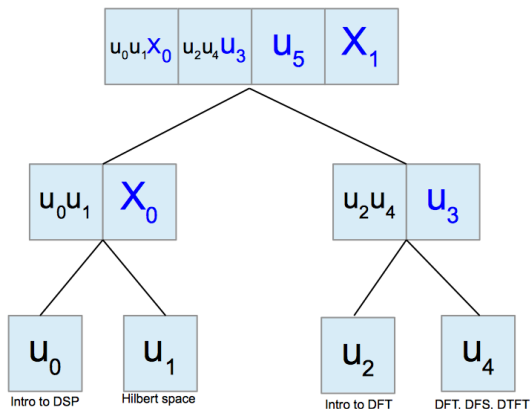
MDR schedule for the DSP MOOC (Spr '13)

Recap for $u_0, u_1, u_2, u_3, u_4, u_5 \dots$



MDR schedule for the DSP MOOC (Spr '13)

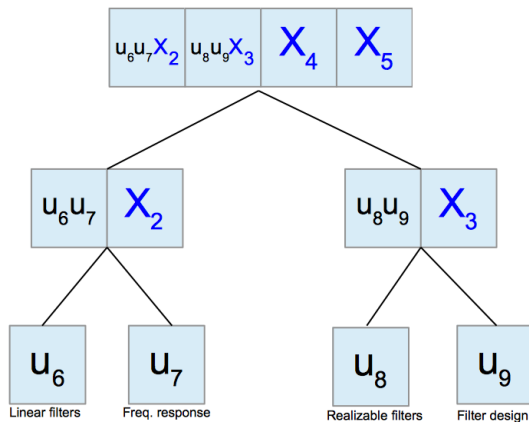
Recap for $u_0, u_1, u_2, u_3, u_4, u_5 \dots$



What to recap in X_0, X_1 ?

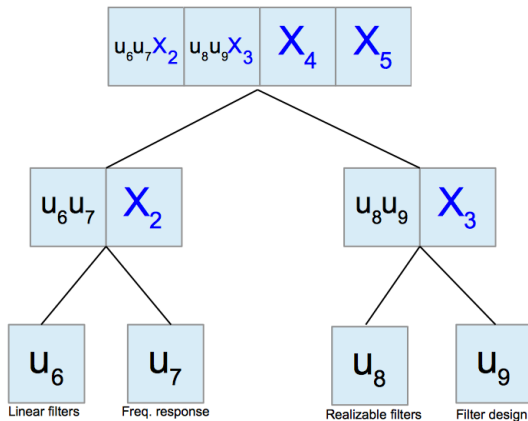
MDR schedule for the DSP MOOC (Spr '13)

Recap for $u_6, u_7, u_8, u_9 \dots$



MDR schedule for the DSP MOOC (Spr '13)

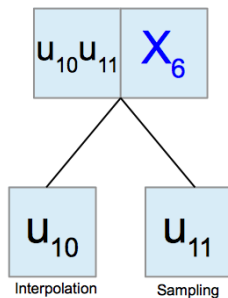
Recap for $u_6, u_7, u_8, u_9 \dots$



What to recap in X_2, X_3, X_4, X_5 ?

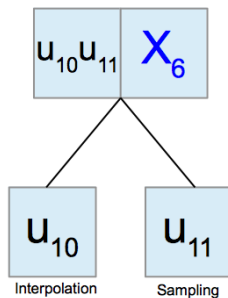
MDR schedule for the DSP MOOC (Spr '13)

Recap for $u_{10}, u_{11} \dots$



MDR schedule for the DSP MOOC (Spr '13)

Recap for $u_{10}, u_{11} \dots$



What to recap in X_6 ?

X_0 recap

- u_0 (intro to DSP) and u_1 (Hilbert space)

Concept filtering

- Signal, basic math, vector space, orthogonality

- Q: Floor (w/ discontinuity), constant functions periodic?
- A: Yes ...

- Q: Properties of sinusoids ...
- A: $W_N^{N/2} = \exp(-j2\pi \frac{N}{2N}) = \exp(-j\pi) = -1$...

- Q: Properties of the inner product ...
- A: Check the axioms of the inner product ...

X_1 recap

X_1 recap

- u_2 (intro to DFT), u_4 (DFT, DFS, DTFT)

Concept filtering

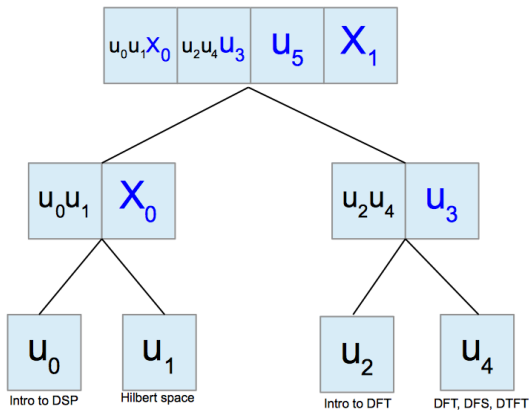
- Fourier analysis, periodicity

- Q: Calculation of DTFT ...
- A: The modulation/convolution/linearity properties of DTFT pairs
- A: DTFT pairs of step/sine/sawtooth functions
- A: Properties of a Hermitian operator

- Q: Discrete Fourier Series vs. Discrete Fourier Transform
- A: Fourier analysis for periodic functions and discrete sequences ...

- Q: A factor $\frac{1}{\sqrt{N}}$ in front of both DFT and IDFT?
- A: Factor $\frac{1}{N}$ considered in the IDFT ...

X_1 recap



X_2, X_3, X_4, X_5 recap

X_2, X_3, X_4, X_5 recap

- u_6 (linear filters), u_7 (frequency response)
- u_8 (realizable filters), u_9 (filter design)

Concept filtering

- Filter

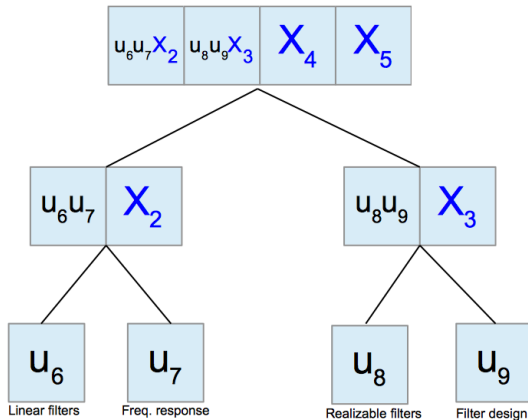
X_2, X_3, X_4, X_5 recap

- Q: Properties of a filter ...
- A: Linearity, stability, causality, linear phase ...

- Q: The transfer function of a filter ...
- A: Analyze the filter regardless of the input ...

- Q: Poles and zeros and filter properties ...
- A: Use ROC inside/outside the unit circle to prove ...
- Q: Moving the poles of a leaky integrator ...
- A: Resulting in a narrow-pass filter ...

X_2, X_3, X_4, X_5 recap



X_6 recap

- u_{10} (interpolation and bandlimited signals)
- u_{11} (sampling and aliasing)

Concept filtering

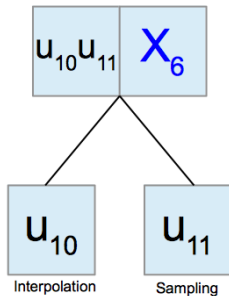
- Sampling, interpolation

- Q: Linear and Lagrange interpolation
- A: Use the samples as bases ...

- Q: Sampling at exactly the Nyquist rate ...
- A: Pointwise effect ...
- Q: How to determine the sampling freq?
- A: BW is known at the receiver side ...

- Q: What is the time duration between two samples ...
- A: $\frac{1}{f}$...

X_6 recap



- $X_0, X_1, X_2, X_3, X_4, X_5, X_6$: Numerical examples

- The recap scheduling problem in education
- The Monotonically Decreasing Recap (MDR) schedule
- MOOC data processing
- The design of MDR for the DSP MOOC
- Conclusions & future work

Conclusions

The recap scheduling problem in education

- Rich ground for modellization

Conclusions

The recap scheduling problem in education

- Rich ground for modellization
- FLR difficult to realize

The recap scheduling problem in education

- Rich ground for modellization
- FLR difficult to realize
- MDR is linear & more practical

The recap scheduling problem in education

- Rich ground for modellization
- FLR difficult to realize
- MDR is linear & more practical

MOOC data processing

- Rich ground for model validation

The recap scheduling problem in education

- Rich ground for modellization
- FLR difficult to realize
- MDR is linear & more practical

MOOC data processing

- Rich ground for model validation
- Filtering forum data is useful

The recap scheduling problem in education

- Rich ground for modellization
- FLR difficult to realize
- MDR is linear & more practical

MOOC data processing

- Rich ground for model validation
- Filtering forum data is useful
- MDR design for the DSP MOOC

Recap schedule modellization

- Dependency between educational units

Recap schedule modellization

- Dependency between educational units
- Flexible & slow students

Recap schedule modellization

- Dependency between educational units
- Flexible & slow students
- Cramming

Recap schedule modellization

- Dependency between educational units
- Flexible & slow students
- Cramming

Model validation in MOOCs

- Filtering algorithms for MOOC data

Recap schedule modellization

- Dependency between educational units
- Flexible & slow students
- Cramming

Model validation in MOOCs

- Filtering algorithms for MOOC data
- Recap schedule to fit students' needs

Recap schedule modellization

- Dependency between educational units
- Flexible & slow students
- Cramming

Model validation in MOOCs

- Filtering algorithms for MOOC data
- Recap schedule to fit students' needs
- Improve MOOCs

Thank you, questions please.

tao.lee@epfl.ch