問題学習に関する新しい理論とそのファジィ論的拡張

2Y-8

矢鳴 虎夫　　興梠 英二

九州工業大学情報科学センター

1．はじめに

一般に教科書の内容を理解しようとする時には，章末などの練習問題でその理解度を把握しようとする．ここでは，その練習問題に準じて，次の３つの事を仮定（前提）におき，問題学習のための新しい理論を提案する．

1)一般に練習問題の内容は教科書を構成する複数個の単元に関連しあっている．

2)問題の難しさと理解度の積は学習量に比例する．

3)難しさは理解度が上がるに従って減少する．

2．理論

2．1　テキスト知識

n個の単元によって構成される教科書を想定し，c個の問題 $\{P_k\}$ k=1,2,…,c が設けられているものとする．そして，これらの問題は単元iと，関連度 $R_i(P_k)$ （i=1,2,…,n　k=1,2,…c）で結びついているものとする．これらを本理論では，テキスト知識と呼ぶことにし，生徒が問題を解く前にあらかじめ，用意されているものとする．具体的には，以下の表に示した通りである．

No. of prob. P_k	NUMBER OF UNITS i 1	2	3	4	5	6	7	8	9	10
1:	.542	.361	.091	.006	.000	.000	.000	.000	.000	.000
2:	.562	.353	.079	.006	.000	.000	.000	.000	.000	.000
3:	.607	.307	.080	.006	.000	.000	.000	.000	.000	.000
4:	.471	.435	.086	.008	.000	.000	.000	.000	.000	.000
5:	.613	.300	.081	.007	.000	.000	.000	.000	.000	.000
6:	.250	.459	.219	.066	.006	.000	.000	.000	.000	.000
7:	.327	.344	.278	.045	.006	.000	.000	.000	.000	.000
8:	.237	.356	.340	.061	.006	.000	.000	.000	.000	.000
9:	.306	.365	.273	.052	.004	.000	.000	.000	.000	.000
10:	.255	.461	.229	.049	.005	.000	.000	.000	.000	.000
11:	.054	.206	.424	.267	.044	.004	.000	.000	.000	.000
12:	.040	.210	.417	.279	.050	.003	.000	.000	.000	.000
13:	.065	.195	.373	.301	.061	.005	.000	.000	.000	.000
14:	.054	.239	.404	.248	.050	.004	.000	.000	.000	.000
15:	.057	.259	.353	.281	.045	.005	.000	.000	.000	.000
16:	.004	.046	.303	.372	.217	.052	.005	.000	.000	.000
17:	.006	.050	.224	.383	.254	.078	.005	.000	.000	.000
18:	.005	.061	.181	.444	.260	.046	.004	.000	.000	.000
19:	.005	.054	.194	.424	.277	.043	.004	.000	.000	.000
20:	.005	.043	.288	.381	.230	.049	.003	.000	.000	.000
21:	.000	.005	.055	.269	.363	.250	.051	.006	.000	.000
22:	.000	.005	.041	.208	.426	.263	.054	.003	.000	.000
23:	.000	.005	.047	.229	.438	.227	.050	.004	.000	.000
24:	.000	.005	.048	.261	.455	.184	.042	.004	.000	.000
25:	.000	.006	.049	.215	.424	.238	.063	.005	.000	.000
26:	.000	.000	.005	.062	.226	.464	.193	.044	.005	.000
27:	.000	.000	.005	.068	.270	.338	.266	.049	.004	.000
28:	.000	.000	.005	.057	.238	.384	.248	.063	.005	.000
29:	.000	.000	.004	.057	.246	.388	.236	.064	.005	.000
30:	.000	.000	.005	.061	.225	.382	.268	.054	.005	.000
31:	.000	.000	.000	.005	.042	.258	.434	.214	.042	.004
32:	.000	.000	.000	.004	.065	.241	.380	.265	.041	.004
33:	.000	.000	.000	.005	.042	.194	.464	.241	.049	.004
34:	.000	.000	.000	.004	.065	.302	.363	.213	.047	.005
35:	.000	.000	.000	.005	.058	.215	.414	.266	.040	.004
36:	.000	.000	.000	.000	.005	.045	.236	.427	.246	.042
37:	.000	.000	.000	.000	.006	.056	.294	.392	.202	.050
38:	.000	.000	.000	.000	.004	.056	.241	.426	.209	.065
39:	.000	.000	.000	.000	.005	.055	.254	.391	.256	.040
40:	.000	.000	.000	.000	.004	.052	.292	.369	.229	.053
41:	.000	.000	.000	.000	.000	.007	.072	.270	.394	.257
42:	.000	.000	.000	.000	.000	.004	.058	.265	.398	.275
43:	.000	.000	.000	.000	.000	.004	.048	.254	.464	.229
44:	.000	.000	.000	.000	.000	.005	.076	.281	.410	.228
45:	.000	.000	.000	.000	.000	.005	.059	.245	.397	.293
46:	.000	.000	.000	.000	.000	.000	.007	.075	.313	.605
47:	.000	.000	.000	.000	.000	.000	.007	.079	.301	.613
48:	.000	.000	.000	.000	.000	.000	.005	.086	.334	.575
49:	.000	.000	.000	.000	.000	.000	.004	.077	.315	.603
50:	.000	.000	.000	.000	.000	.000	.006	.068	.347	.579

表１．関連度表

ただし，

$$\sum_{i=1}^{n} R_i(P_k)=1 \qquad (1)$$

$(R_i(P_k)\geq 0$ i=1,2,…,n　k=1,2,…c)　である．

2．2　学習方法

学生は１セションの学習でm個の問題を学習するものとし，学習結果は担当教師に渡され採点されるものとする．以後t番目のセションを'学習ステップt'ということにし，ある学生のtで学習するm個の問題を $\{P_1(t), P_2(t), \cdots, P_m(t)\}$ で記すことにする．（この $P_r(t)$ といった量は，本来学生個人を識別するための添え字が必要であるが，この理論は一学生を対象としたものであるため，この添え字はいっさい省略する．以後で定義する量も又同じである）

2．3　ファジィ理論の導入

テキスト知識をこと細かに記述するのは，なかなか難しい問題である．例えば，表１において，問題１が単元１に0.542関与していると決定することは，一般には，困難である．

そこで，ここでは，関連度の曖昧さというのを考慮にいれたファジィ理論への展開を考え，この問題に対処する．つまり関連度表の $R_i(P_k)$ の値に，曖昧さが　あると仮定して，メンバーシップ関数を導入するわけである．そこで理論に曖昧さという幅をもたせて，それにもとづくファジィ理論を導入し，ファジィ的学習理論を定義する．

テキスト知識について，導入された理論を次に示す．

関連度を

$$R_i(P_k)=\mu_{R_i(P_k)}(x)/x \qquad (2)$$

で表わす．但し $x\in X=[0,1]$ とする．

以下に定義する種々の概念量もまた，この範囲内であるとする．

2．4　問題の難しさの決定

単元iのステップtでの難しさを $\mathbb{D}_i(t)$ で記すことにする．$\mathbb{D}_i(t)$ はtとともに1～0の範囲で変化するものとする．特に初期値 $\mathbb{D}_i(0)$ はその学生の初期能力もしくは，前提知識に対応するもので，その単元iに関する知識をもたないときは $\mathbb{D}_i(0)=1$ にする．

このとき問題 $P_r(t)$ のステップtにおける難しさは，

$$\mathbb{D}(P_r(t),t)=\sum_{i=1}^{n}\mathbb{D}_i(t)R_i'(P_r(t)) \qquad (3)$$

で与えられる．ただし，R_i' は規格化関連度で次のように定義される．

$$R_i'(P_r(t))=R_i(P_r(t))/A_i\sum_k R_i(P_k) \qquad (4)$$

A_i：単元iを完全に理解するのに必要な最小問題数の単元iに関連する全問題数に対する割合．

$$0<A_i\leq 1 \qquad (5)$$

$\Sigma R_i(P_k)$：単元iに関連する全ての問題の関連度の和　(6)

ただし，ここで定義した $\mathbb{D}$ も曖昧さを持ったファジィのメンバーシップ関数で，表されているとする．

この R_i' を設けた理由は，一般に教師はその単元に関連したすべての問題を与えなくても数問に対してある程度の得点をとれば，その単元は理解したと見なすのが普通である．このことを理論化するために A_i というパラメータを設けた．例えば，単元iに関連する問題が全部で１０問（$P_{i_1}, P_{i_2}, \cdots, P_{i_{10}}$）　であって，$A_i=0.5$ と担当教師が指定

A New Theory Concerning the Study for exercise Problems and it's Fuzzy theoretical expansion.

Eiji Kourogi

Information Science Center,Kyushu Institute of Technology

した場合には，この１０問のうち５問満点とればこの単元ｉの理解度はほぼ完全と見なす．従ってステップｔのｒ番目に提示される問題$P_r(t)$は，一般に学生が選択した単元群｛$i_1, i_2, \cdots, i_g$｝のうち最も容易な問題を提示するという一般的学習原理に従えば，

$$P_r(t)=P_k \mid \mathrm{mini}\,[\,\mathbb{D}(P_k,t), P_k \in \{\mathrm{all}\ P_k \mid i=i_1, i_2, \cdots, i_g\}\,] \qquad (７)$$

と定義できる．しかし，$\mathbb{D}(P_k,t)$ はメンバーシップで表されているので，最小値を計算する場合はファジィメンバーシップの頂点（PEEK）の値をもって比較するものとする．

２．４　種々の概念の量

理解度，学習量，努力量といった概念の定量化を試みる．ステップｔで学生に提示され，学習された問題群｛$P_r(t)$，$r=1,2,\cdots,m$｝は担当教師によって評価される．

$$\{M(P_r(t)),\ r=1,2,\cdots,m\} \qquad M\text{は得点．} \qquad (８)$$

この時これらの評価点はその問題に対するその学生の'理解度'と定義する．

従ってステップｔでのｍ個の問題をといた時の単元ｉに関する理解度の増分$\Delta\mathbb{U}_i(t,m)$ は，一度理解したことは忘れないという仮定と理解度の変化の範囲は０～１という仮定を設ければ，

$$\Delta\mathbb{U}_i(t,m)=\sum_{r=1}^{m}\Delta\mathbb{U}_i(P_r(t)) \qquad (９)$$

$$=\sum_{r=1}^{m}M(P_r(t))\cdot R_i'(P_r(t)) \qquad (１０)$$

で与えられる．一方，仮定より，単元の理解度の増加にともなって，単元の難しさは次第に小さくなる，と考える．そのためこの性質を数式的に取り入れるため，各単元において１ステップ前の難しさを土台にして学習し，そして理解し，難しさの減少をきたすと考えた．つまり

$$\mathbb{D}_i(t)=\mathbb{D}_i(t-1)\,(1-\Delta\mathbb{U}_i(t,m)) \qquad (１１)$$

である．

次に学習量について考える．これも前提より，問題Ｐｒをその時の難しさＤで学習して得点Ｍ(Ｐｒ)を得るときの学習量Ｓは次の式で定義する．

$$\Delta\mathbb{S}_i(t,m)=\beta\cdot\mathbb{D}_i(t-1)\,\Delta\mathbb{U}_i(t,m) \qquad (１２)$$

$\Delta\mathbb{S}_i$ の変化の範囲も$\mathbb{D}_i$ や$\Delta\mathbb{U}_i$ と同様に０～１と規格化すれば，$\beta=1$となって，

$$\Delta\mathbb{S}_i(t,m)=\mathbb{D}_i(t-1)\,\Delta\mathbb{U}_i(t,m) \qquad (１３)$$

となる．つまり，単元ごとの学習量は理解度と難しさの積として定義する．

従ってｔステップまでの全学習量は

$$\mathbb{S}_i(t,m)=\sum_{e=1}^{t}\Delta\mathbb{S}_i(e,m) \qquad (１４)$$

で与えられる．

また，$\mathbb{S}_i(t,m)$ の$t=t_{max}$（単元ｉに関連する問題の全てが学習されて提示する問題がなくなるまでのステップ）のとき最大値をもって単元ｉのｍ個／ステップ方式での学習容量$\mathbb{S}_{ic}$ということにすれば，

$$\mathbb{S}_{ic}=\max\{\mathbb{S}_i(t,m)\} \qquad (１５)$$

で与えられる．

さらに，個々の学生の努力量として学習容量に対してどれほど学習したかという尺度を与えれば，

$$\mathbb{E}_i(t,m)=\mathbb{S}_i(t,m)/\mathbb{S}_{ic}(m) \qquad (１６)$$

と定義することができる．

２．５　単元群への拡張

次に今までの学習単元に基づいた理論を単元に拡張すると，理解度，学習量などの一連の量を平均値として次のような式を導くことができる．いま，学生が学習しようと選択する単元群を$G=\{i_1, i_2, \cdots, i_g\}$とすると，

$$\Delta\mathbb{U}_G(t,m)=\sum_{j=1}^{g}\frac{1}{g}\Delta\mathbb{U}_{i_j}(t,m) \qquad (１７)$$

$$\mathbb{U}_G(t,m)=\left(\sum_{e=1}^{t}\mathbb{U}_G(e,m)\right) \qquad (１８)$$

$$\mathbb{D}_G(t)=\sum_{j=1}^{g}\frac{1}{g}\mathbb{D}_{i_j}(t) \qquad (１９)$$

$$\Delta\mathbb{S}_G(t,m)=\sum_{j=1}^{g}\frac{1}{g}\Delta\mathbb{S}_{i_j}(t,m) \qquad (２０)$$

$$\mathbb{S}_G(t,m)=\sum_{j=1}^{g}\frac{1}{g}\mathbb{S}_{i_j}(e,m) \qquad (２１)$$

$$\mathbb{S}_{Gc}(m)=\sum_{j=1}^{g}\frac{1}{g}\mathbb{S}_{i_c}(m) \qquad (２２)$$

$$\mathbb{E}_G(t,m)=\sum_{j=1}^{g}\frac{1}{g}\mathbb{E}_{i_j}(t,m) \qquad (２３)$$

一方(12)式から，

$$\mathbb{D}_i(t-1)=\Delta\mathbb{S}_i(t,m)/\Delta\mathbb{U}_i(t,m) \qquad (２４)$$

となるので，$\mathbb{D}_i(t-1)$ は学習効率を表わすことにもなっている．この考え方を拡張すれば，等価学習効率$\mathbb{D}_{GEQ}(t,m)$を

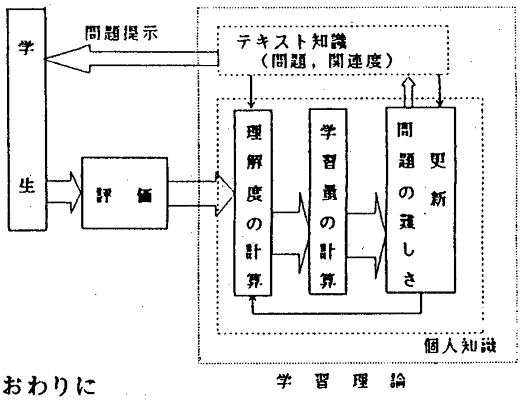
$$\mathbb{D}_{GEQ}(t,m)=\mathbb{S}_G(t,m)/\mathbb{U}_G(t,m) \qquad (２５)$$

として定義できる．

３．システム化

３．１　概要

以下に示すようなシステムとなる．



４．おわりに

関連度表の$R_i(P_k)$を問題ごとに与えることは，実際の教科書のインターフェイスにおいて，まだまだ問題点がある．将来この$R_i(P_k)$の理論的考え方，またはその同定法について研究する余地がある．

また，この理論は，教科書をいくつかの単元に分けて考えているが，この単元については，概念であってもよいと考える．