

厚生労働科学研究費補助金

政策科学総合研究事業（統計情報総合研究）

「疾病、傷害及び死因統計分類」の変更が
わが国の厚生統計に与える影響に関する研究

(H27－統計－一般－002)

平成28年度 総括・分担研究報告書

研究代表者 緒方 裕光

平成 29（2017）年 5 月

目 次

I. 研究報告書抄録

「疾病、傷害及び死因統計分類」の変更がわが国の厚生統計に与える影響に関する研究……1

緒方裕光、水島洋、佐藤洋子

II. 資料

1. 分類変更が患者調査、人口動態統計に与える影響……7

2. 患者調査に関する影響解析……11

3. 人口動態に関する影響解析……47

厚生労働科学研究費補助金（政策科学総合研究事業（統計情報総合研究））
総括研究報告書

「疾病、傷害及び死因統計分類」の変更がわが国の厚生統計に与える影響
に関する研究

研究代表者 緒方 裕光 国立保健医療科学院 研究情報支援研究センター長

研究要旨

本研究では、ICD など疾病や死因分類の変更が厚生統計に与える影響を定量的に把握することを目的として、分類変更前後の変化を時系列的かつ統計学的に推定するためのモデル及び方法論を検討・提案し、この方法に基づき、分類変更が人口動態統計や患者調査などへ与える影響を定量的に評価する。

前年度においては、分類変更の基本的パターンに基づいて統計的モデルの構築を行い、シミュレーションや実データの解析を通じて分類変更時の不連続の検出、モデルの評価などを行い、本モデルの適用可能性を示した。この結果に基づき本年度（平成 28 年度）においては、患者調査および人口動態統計について、本モデルを用いて ICD コードの 2003 年版から 2013 年版への分類変更パターンを可視化したうえで、その影響を定量的に評価した。

本研究で提示したモデルを用いることにより、変更後の統計データの変動予測が可能である。これらの予測は、とくに分類変更により大きな影響を受ける疾患や死因に関しては、真の増減を見出す際に重要な情報となる。

研究分担者

水島 洋 国立保健医療科学院
研究情報支援研究センター
上席主任研究官
佐藤洋子 国立保健医療科学院
研究情報支援研究センター
研究員

の整合性、国際比較可能性、正確性などの低下の原因となる。厚生労働統計の有用性を高めるためには、このような分類変更に伴う影響を合理的な方法で評価し、将来推計やデータ比較の際に適切な補正を行う必要がある。

分類変更の影響に関する研究は、欧米を中心に多数の研究例があり、これらは主に 2 種類に分けられる。第 1 は、変更前の一定期間にわたって同一のケースを変更前後 2 種類の分類方法でコーディングする方法（ダブルコーディング）であり、第 2 は、類似のケースを変更前後それぞれの分類でコーディングする方法である。前者は直接的に変更前後の分類の関係を知らることができる。後者は、クロス表を作成することに

A. 研究目的

ICD（疾病及び関連保健問題の国際統計分類）の変更や部分的な修正などによる分類方法の変更によって、分類変更の前後で特定の分類項目に関して見かけ上の統計数値の増減（不連続）が生ずる。このことは、我が国において関連する統計データの相互

よって間接的に変更前後の分類の関係を知ることができる。しかし、これらの方法は、コストや労力がかかること、特定の国にのみ応用できること、大きな変更時の1、2年程度の間には有効な方法であることなどの欠点がある。これに対して、長期間にわたる時系列データを利用した定量的方法もいくつか提示されているものの、モデルは統一的ではなく、特定の死因に限られるなど、汎用性という点では十分ではない。

本研究では、分類変更の厚生労働統計への影響を定量的に把握することを目的として、分類変更前後の変化を時系列的かつ統計学的に推定するためのモデル及び方法論を検討・提案し、この方法に基づき、分類変更が人口動態統計や患者調査などへ与える影響を定量的に評価する。本研究は、主にモデル構築とその妥当性の検証、およびモデルによる変更の影響評価からなる。2年計画の初年度（平成27年度）には、主にモデルの構築及びその適用可能性の検討を行う。平成28年度にはモデルの妥当性の検証を行った上で、分類変更が実際に患者調査や人口動態統計へ与える影響の定量的評価を行う。

B. 研究方法

本研究では、統計学的な理論に基づきモデルの構築を試みる。本研究におけるモデルの基本的考え方は以下のとおりである。

ある時点で分類基準をAからBに変更したと仮定し、分類Aに基づく結果をX、分類Bに基づく結果をY'とする。分類基準変更前に分類Bを用いたと仮定した場合の結果をY、一方、分類基準変更後も分類Aを用いたと仮定した場合の結果をX'とおく。Y'とXの差が分類変更前後の結果の見かけ上の差であるが、この差には時間経過後のXとX'の差が含まれるので、真の分類変更による結果の影響は以下のよう

にモデル化できる。

（真の分類変更の影響）

$$= (Y' - X) - (X' - X) = (Y' - X')$$

本研究では、分類変更前の変化を時系列的に見ることによりXからX'を統計学的に推定し、分類変更時の各分類コードの対応に関する情報を利用してX'とY'の関係を求めることにより真の分類変更の影響を定量的に把握する方法を提示する。

全体の研究としては、平成27年度には、統計的モデルの構築を行い、分類変更時の不連続の検出、モデルの適用性評価などを行う。平成28年度には、さらにモデルの一般化およびシミュレーション等による妥当性を評価したうえで、分類変更が患者調査などに与える影響を定量的に評価する。

平成27年度における具体的な研究方法は以下の通りである。

1) 統計的モデルの構築

モデル化にあたっては、いくつかの基本的パターンを考慮して、想定される変数間の関係性を組み込んだ単純モデルを仮定する。この際、統計的なバラツキを考慮しつつ、分類変更がない場合のこれらの時間的変動をモデル化する。

2) モデルの適用性の検証（シミュレーション）

統計学的にある程度のバラツキと傾向性を持つ仮想データを生成し、基本的な分類変更パターンについて、上記モデルの適用可能性を検証する。この際、時間的な変動についてはいくつかの時系列モデルの適用を試みる。

3) 実際のデータへの適用

モデルの実際のデータへの適用の際の課題を検討するため、予備的検討として、入手可能なデータを用いて、分類変更の前後数年間の時系列データに上記モデルの適用を行い、実際の分類変更の影響について統計学的評価を行う。

4) 患者調査や人口動態統計への影響

平成 28 年度には、方法論をさらに一般化するために、より複雑なパターンについてシミュレーションを行い、モデルの妥当性を評価する。また、本研究で提示したモデルを用いて、患者調査や人口動態統計データ等における分類変更の影響を定量的に評価・予測する。

(倫理面への配慮)

本研究では個人情報とは取り扱わない。ただし、患者調査、人口動態統計などの調査情報の目的外利用に関しては、該当する担当部署に必要な申請を行い、承認を得る。

C. 研究結果

平成 27 年度には、統計的モデルの構築を行い、分類変更時の不連続の検出、モデルの適用性評価などを行った。主な結果は以下の通りである。

1) モデルの構築

基本的な分類変更のパターン（1 つの分類コードが複数の分類コードに分かれる場合、複数の分類コードが 1 つの分類コードにまとめられる場合、1 つの分類コードの一部が別の分類コードに移動する場合など）を想定し、必要なパラメータ（時間的変動、変更パターンに応じた分類変更の影響、誤差など）を組み込んだモデルを構築した。

2) シミュレーションデータによるモデルの適用性の検証

統計学的にある程度のバラツキと傾向性を持つ仮想データを生成し、上記モデルの適用可能性を検証した。この際、時間的な変動についてはいくつかの時系列モデルの適用を試みた。その結果、基本的な分類変更パターンについては、十分に適用可能であることを確認した。

3) 実際のデータへの適用

2003 年に分類変更が行われた難治性疾

患の例について、2001 年から 2008 年の集計データを用いて上記モデルの適用を行った。この際、2003 年以前のデータに基づき 2004 年以降の予測を行い、実際のデータと比較することによって予測精度を確認した。その結果、分類変更パターンによっては大きな誤差を生じる場合があることがわかった。

平成 28 年度には、本研究で示した方法の実際の厚生統計への適用を目的として、以下の研究を行った。

4) 患者調査および人口動態統計への適用

「疾病、傷害及び死因の統計分類（ICD-10（2003 年版）準拠）」は、統計法（平成 19 年法律第 53 号）の規定に基づき、統計基準として、ICD-10（2013 年版）に準拠する改正が行われた。改正された ICD-10 は、平成 28 年 1 月 1 日から施行し、同日以後に作成する公的統計の表示に適用されることとなっている。これに伴い、平成 29 年患者調査及び人口動態調査は改正後の ICD-10（2013 年版）で表示することとなるが、その際に統計データについてどのような影響が生じるかを、平成 26 年患者調査および人口動態調査の情報を用いてこの改正（分類変更）の影響を評価した。患者調査については全コード、人口動態調査については、変更の影響が大きいと思われる ICD-10 コード；C80（部位の明示されない悪性新生物）～C97（独立した(原発性)多部位の悪性新生物）に関して、それぞれ変更パターンを可視化し、定量的な評価（平成 29 年度の予測）を行った。その結果、コードによっては統計データが大きく影響を受けることが分かった。

D. 考察

モデル構築の最初の段階としては現実の状況を記号的に明示化することが重要であ

り、本研究で提示したモデルは基本的な分類変更パターンについては十分対応可能であると思われる。

本研究の特色は、分類変更の影響を定量的に評価するための一般的モデル（および方法論）を提示する点にある。本研究で提示したモデル（または方法）は患者調査や人口動態統計等における分類変更の影響を定量的に評価する際には非常に有用であることが分かった。

さらに、分類変更の際には、コーディングのエラーに起因するデータの変動もあり、実際のデータへモデルを適用することによりそれらの変動要因を統計学的に考察する際の根拠にもなりうる。

E. 結論

分類変更の基本的パターンに基づき、その影響を評価するためのモデル構築を行った。さらにシミュレーションや実データの解析を通じてモデルの適用可能性を検討した。本モデルを患者調査および人口動態統計データに適用した結果、分類変更後の予測や統計的評価において有用であることを示した。

F. 健康危険情報

なし。

G. 研究発表

なし。

H. 知的財産権の出願・登録状況

なし。

II. 資料

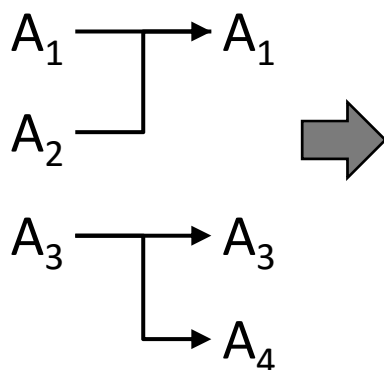
分類変更が患者調査、人口動態統計 に与える影響

佐藤洋子、水島 洋、緒方裕光

国立保健医療科学院
研究情報支援研究センター

単純な分類変更パターン

変更前 変更後



変更前後の分類コードの対応表

変更前	変更後	備考
A ₂	A ₁	A ₂ 削除
A ₃ の一部	A ₄	新設

影響の評価に必要なデータ

分類コード	変更前の該当者数	変更後の該当者数
A ₁	N ₁	N ₁ + N ₂
A ₂	N ₂	/
A ₃	N ₃	N ₃ - N ₄
A ₄	/	N ₄

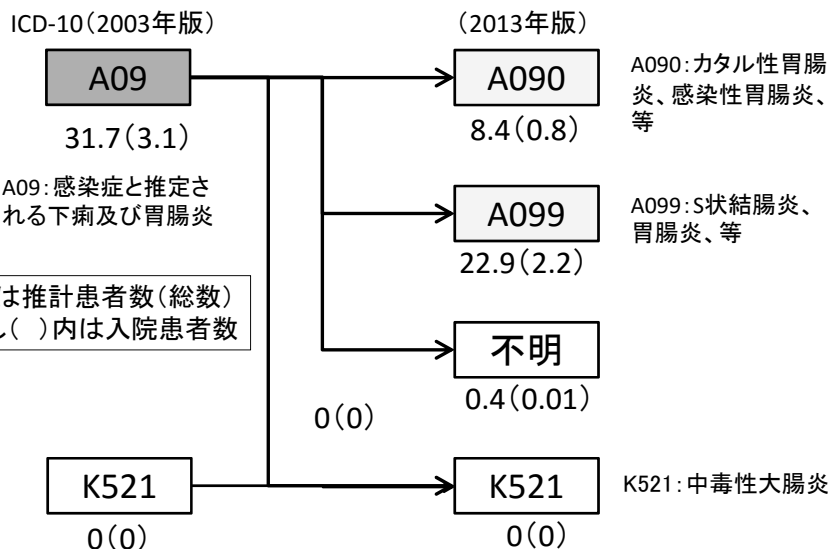
※実際はより複雑なパターン。
ただし、基本パターンの組み合わせ

患者調査で分類変更のあったコードの数 ICD-10(2003年版)から(2013年版)

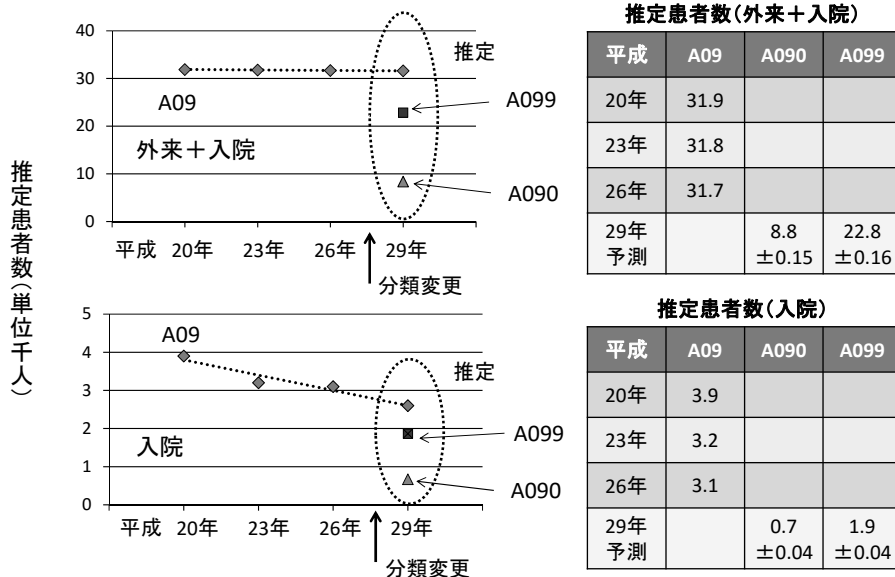
ICD-10 2003コードのうち、2013年版で分類変更分類変更を伴うコードは228個。このうち、平成26年の患者調査データに記載されているコードは189個

推計患者数	外来	入院	退院	入院+ 外来
個票なし	28	31	7	14
100人未満(0>=0,<100)	78	91	102	68
1000人未満(100>=,1000<)	59	45	54	69
10000人未満(1000>=,10000<)	19	20	24	31
10万人未満(10000>=,100000<)	5	2	2	6
10万人以上(100000>=)	0	0	0	1
計	189	189	189	189

分類変更の可視化と推定患者数 (平成26年のA09の例)



患者調査データに対する分類変更の影響 (コードA09の例)



人口動態統計の検討コードと人数

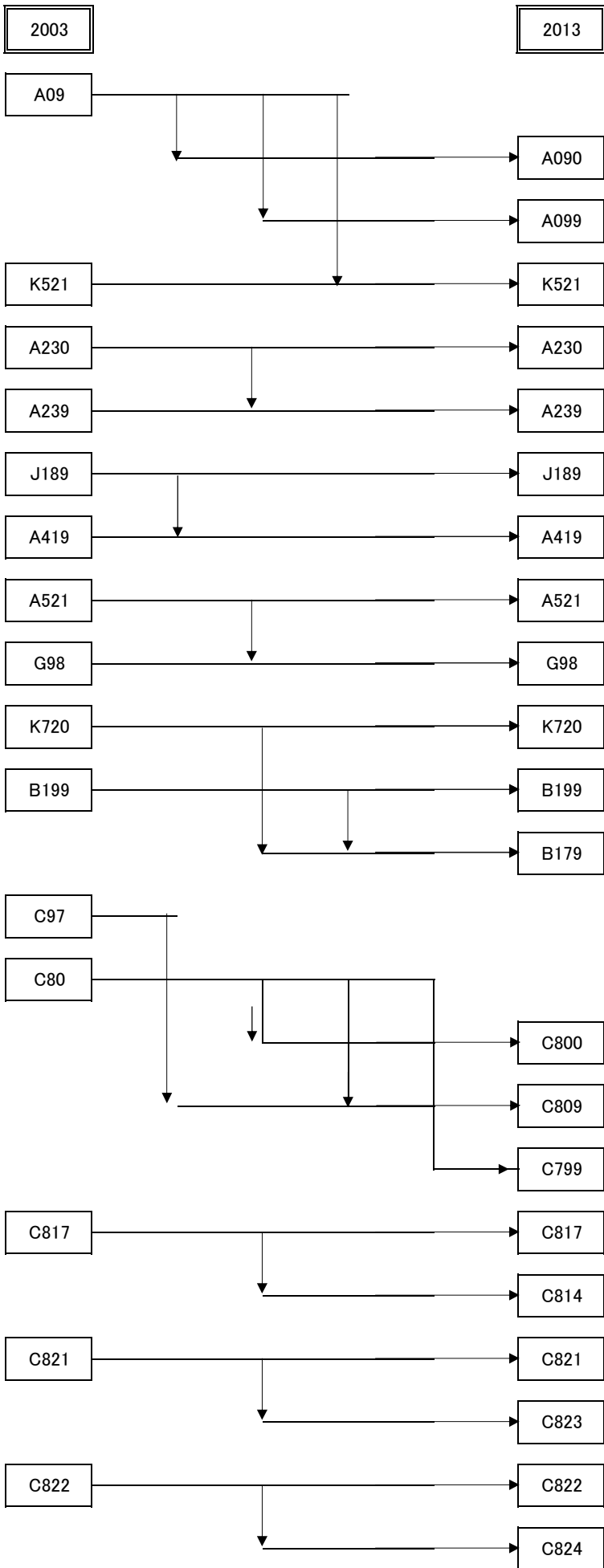
平成26年度人口動態調査に関して、C80(部位の明示されない悪性新生物)～C97(独立した(原発性)多部位の悪性新生物)について患者調査と同様に検討

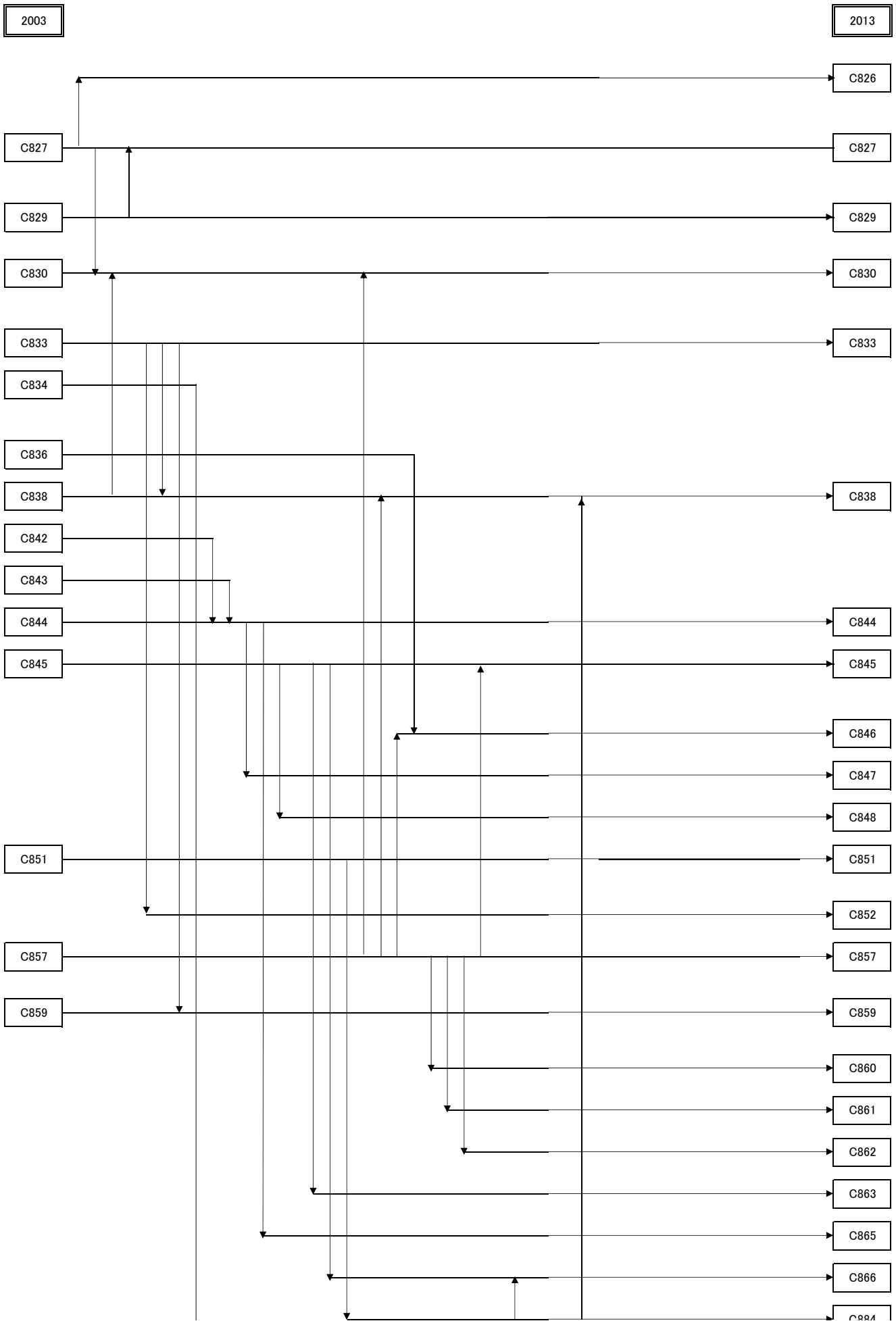
コード	人数	コード	人数	コード	人数	コード	人数	コード	人数	コード	人数
C817	1	C857	32	C927	281	C859	9703	C880	149	C930	36
C944	1	C925	32	C959	403	C811	5	C882	2	C940	10
C947	1	C902	53	C921	417	C812	2	C901	60	C942	11
C836	2	C834	67	C950	470	C819	170	C910	705	C951	27
C917	2	C844	106	C97	598	C835	40	C911	256	C969	13
C827	4	C838	111	C833	799	C837	34	C912	2		
C967	8	C851	112	C915	989	C839	13	C914	8		
C913	9	C829	115	C900	3949	C840	42	C919	25		
C931	10	C929	116	C920	4332	C841	8	C923	5		
C961	31	C845	175	C80	4733	C850	2	C924	105		

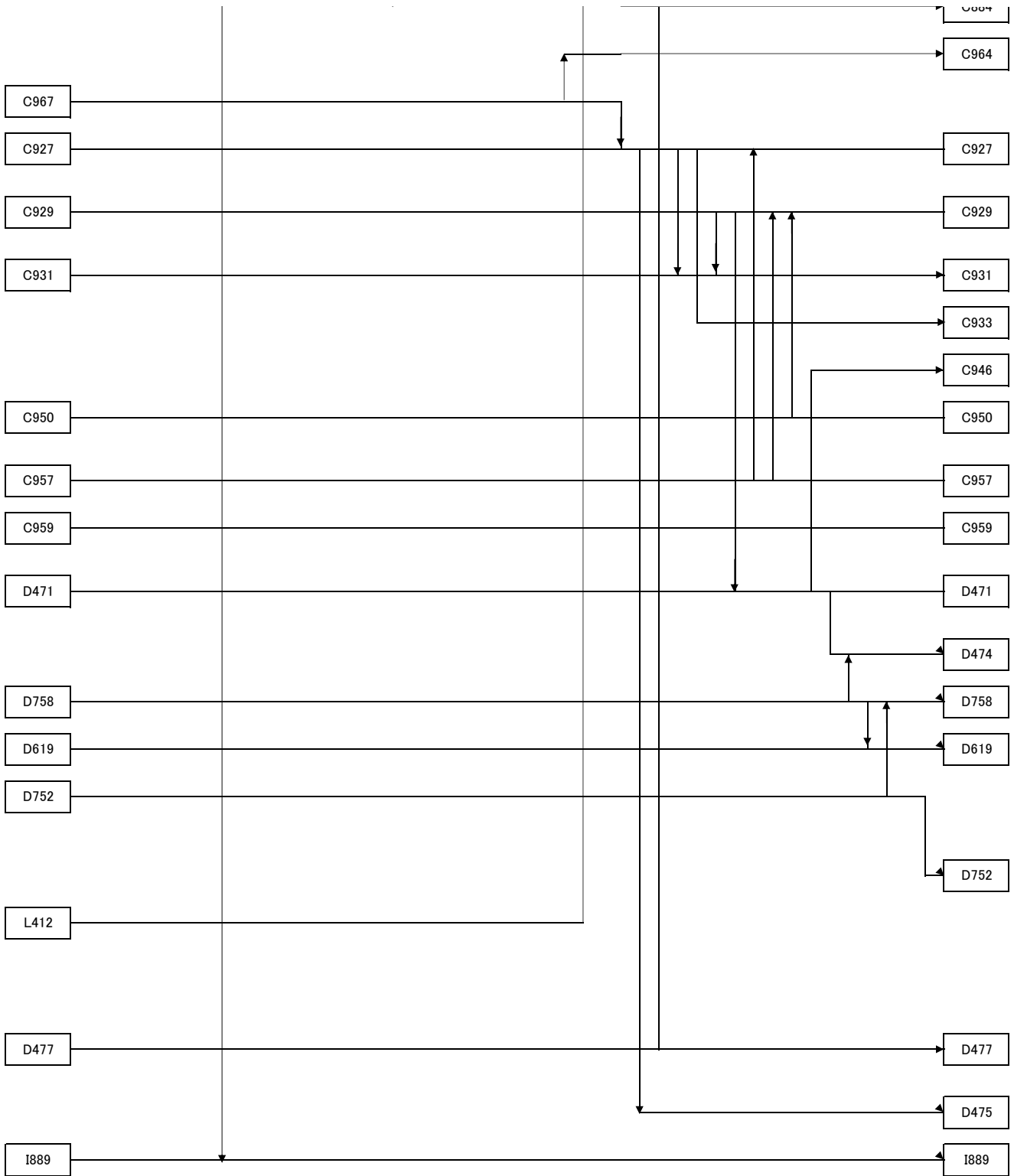
上記のうち、色の濃い部分は、2003年版から2013年版でコードに変更なし

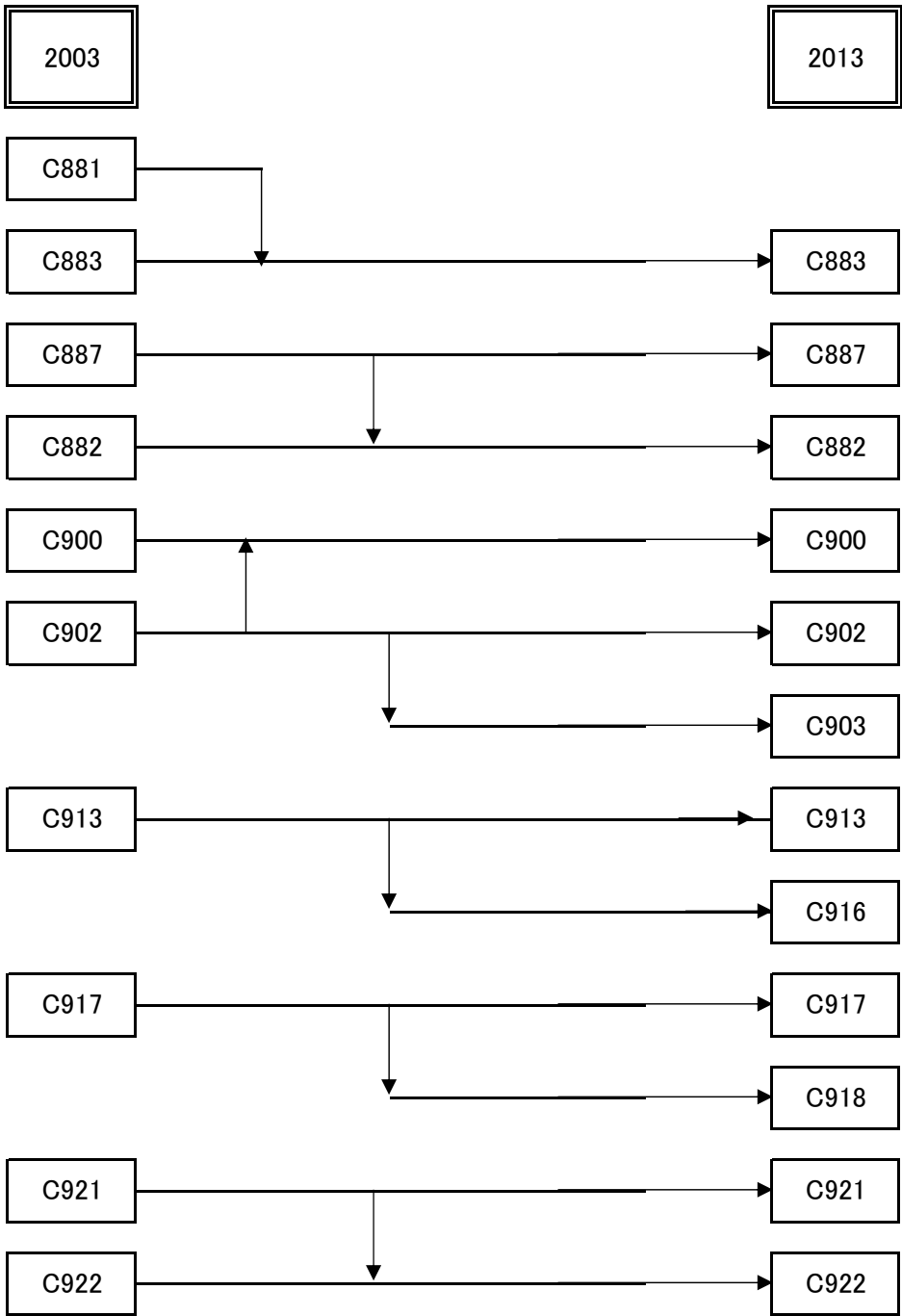
まとめ

- 分類変更の基本的パターンに基づきモデルの構築を行った。
- シミュレーションおよび実際のデータにより方法の適用可能性を示した。
- 人口動態統計、患者調査などに適用した。
- 変更前後の分類コードの対応が合理的であれば時系列的予測や影響の評価が可能。



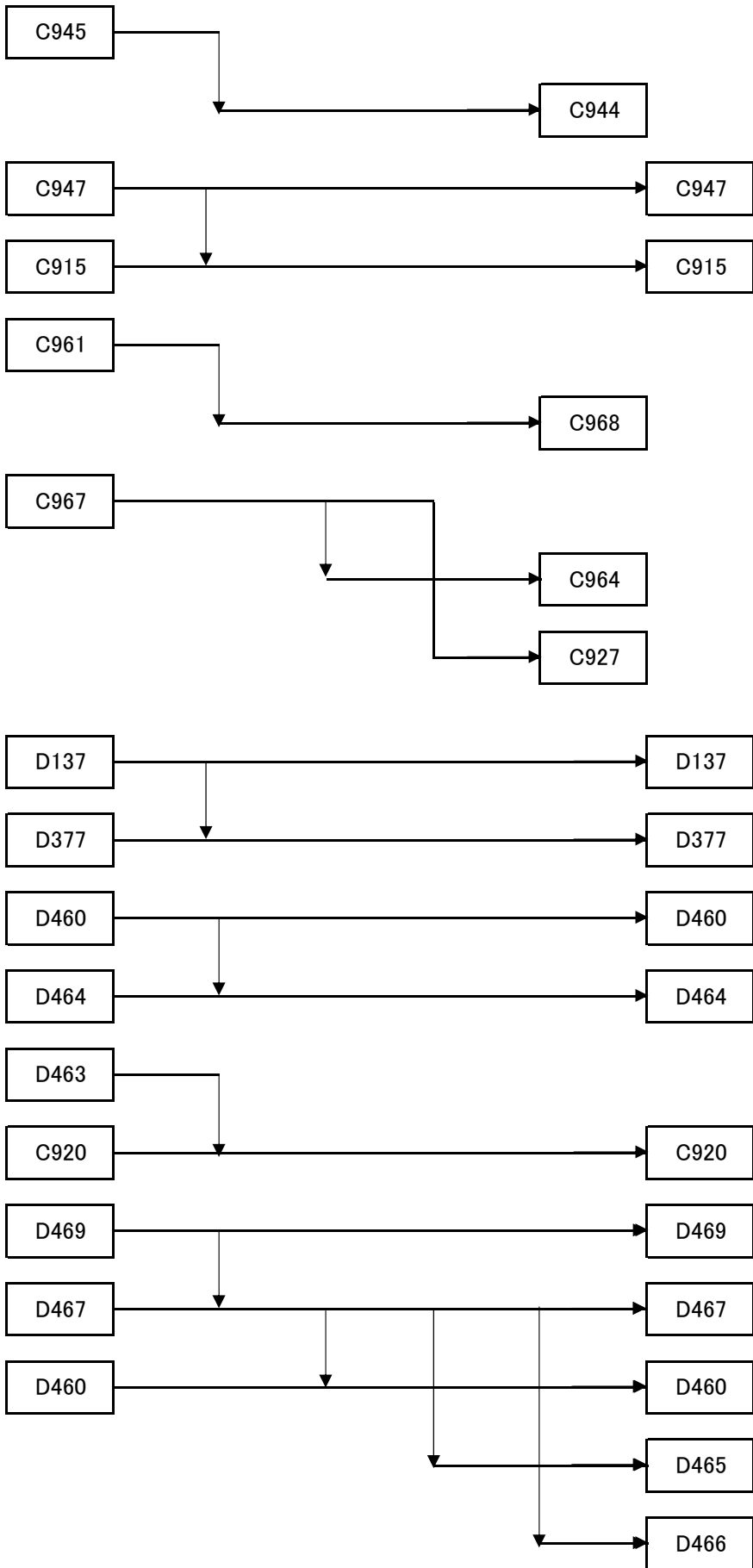


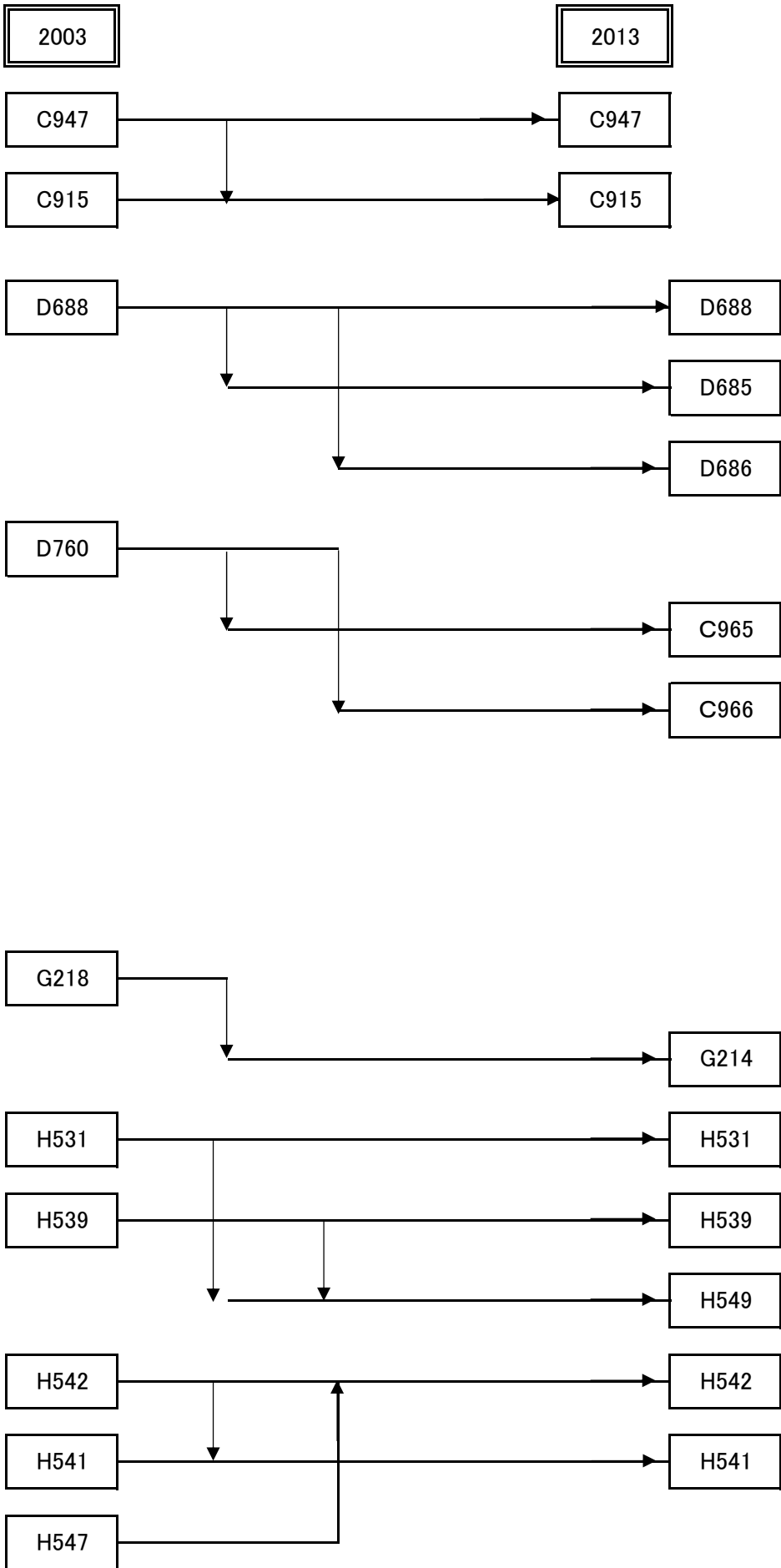


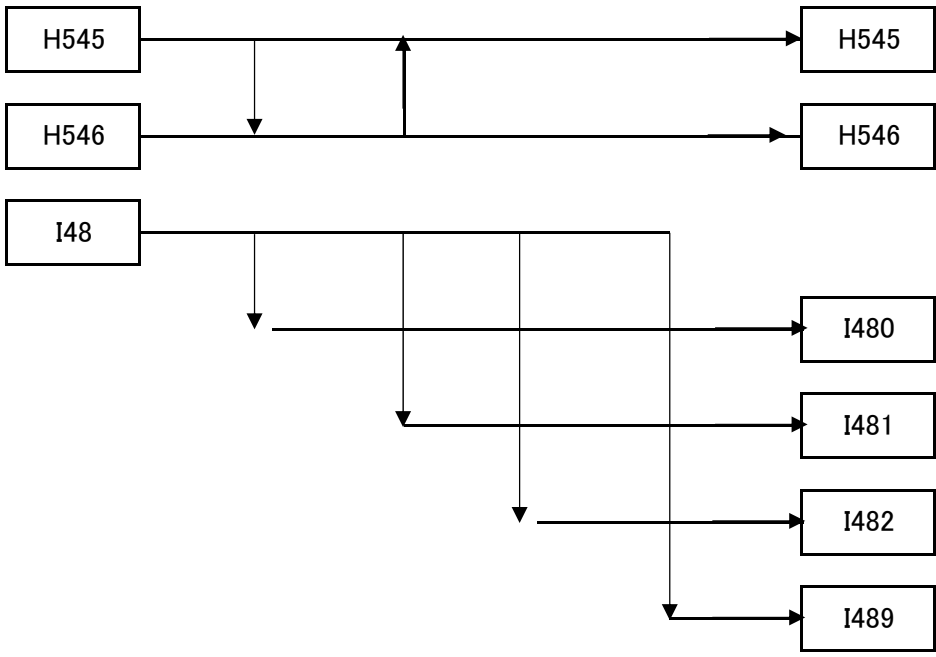


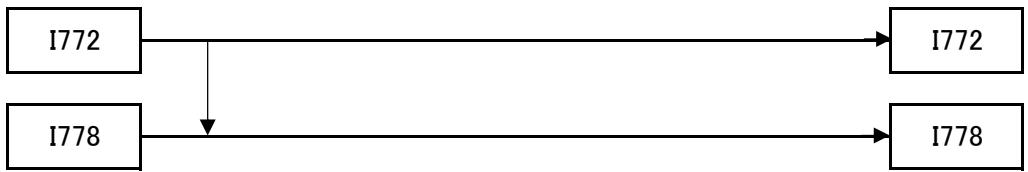
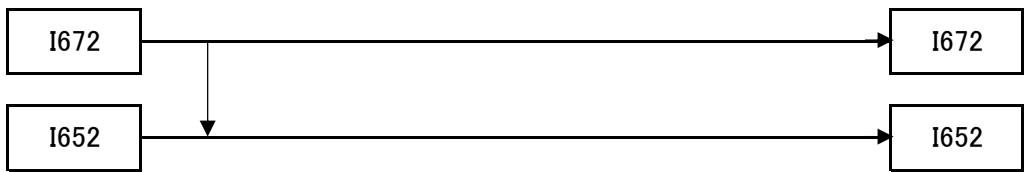
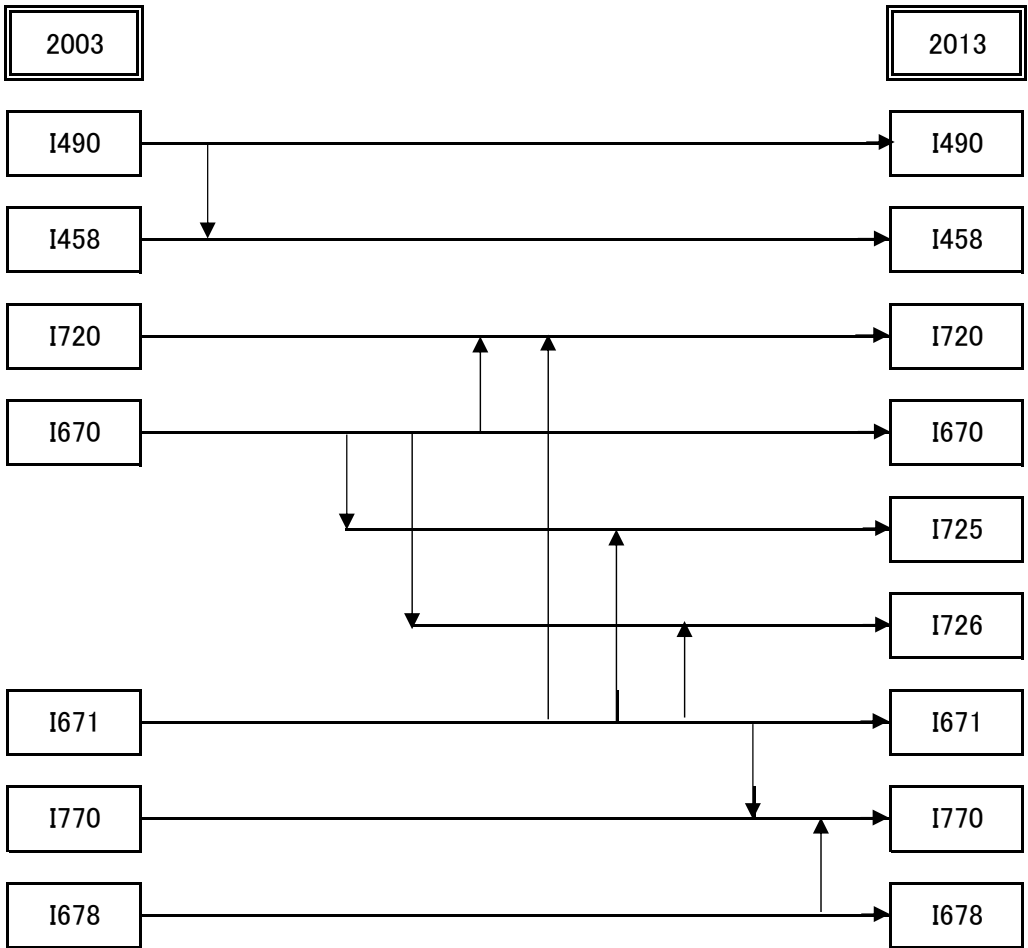
2003

2013



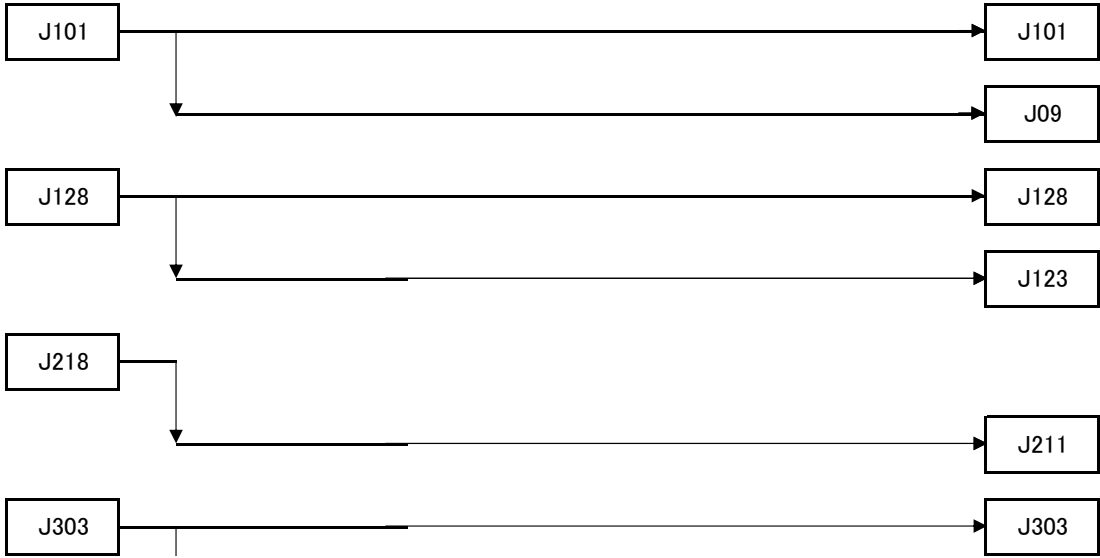
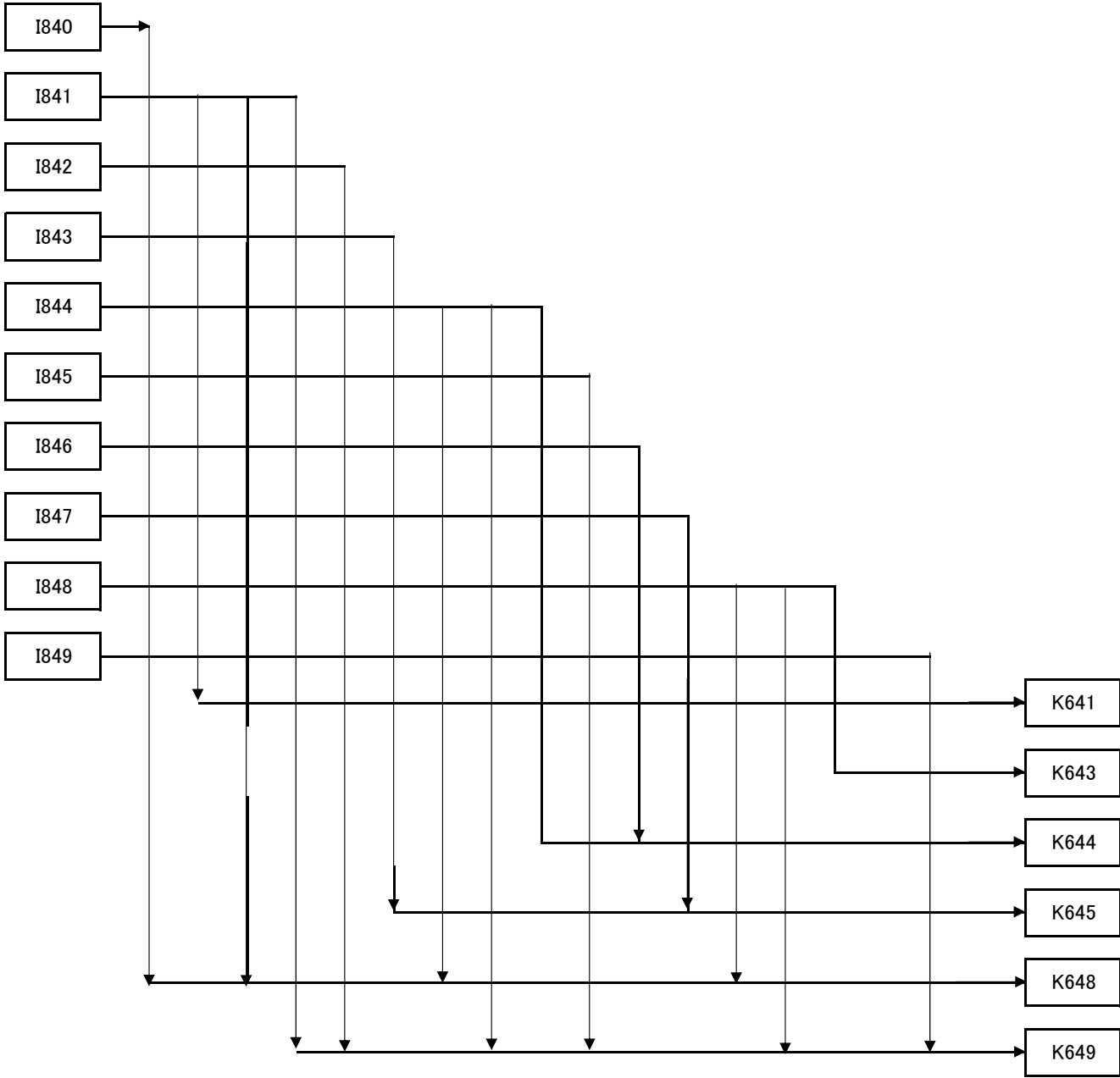


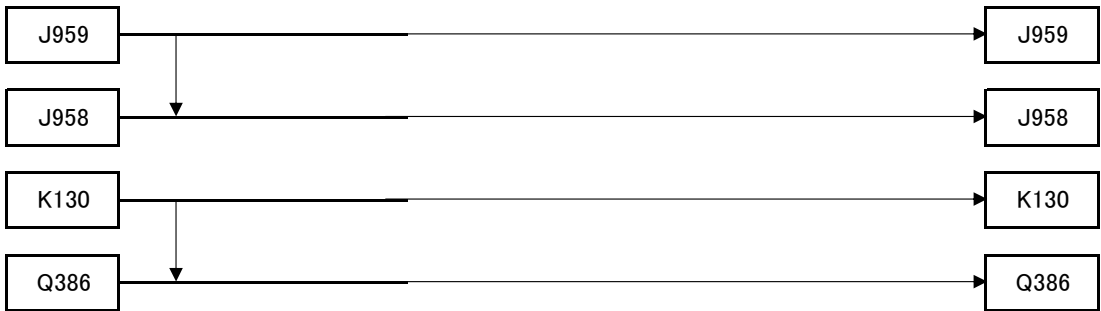
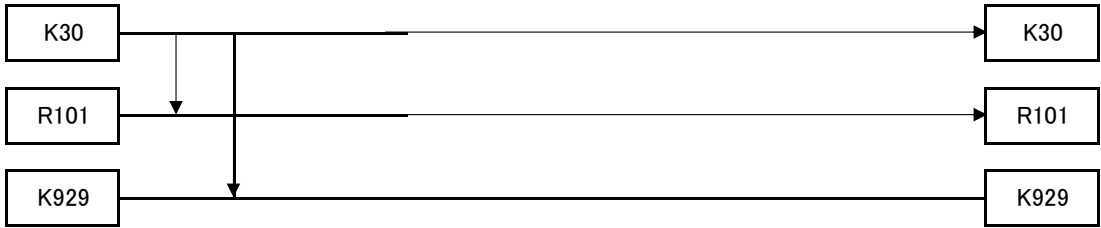
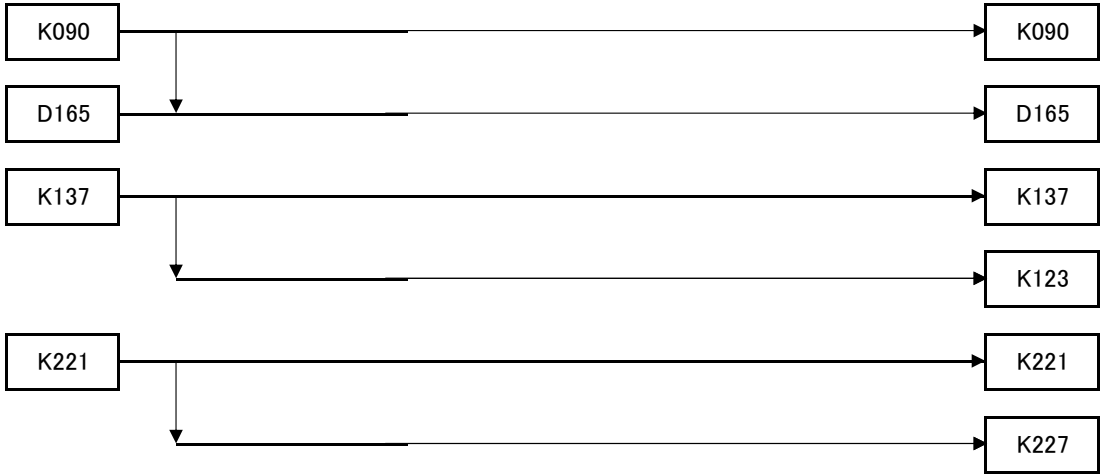
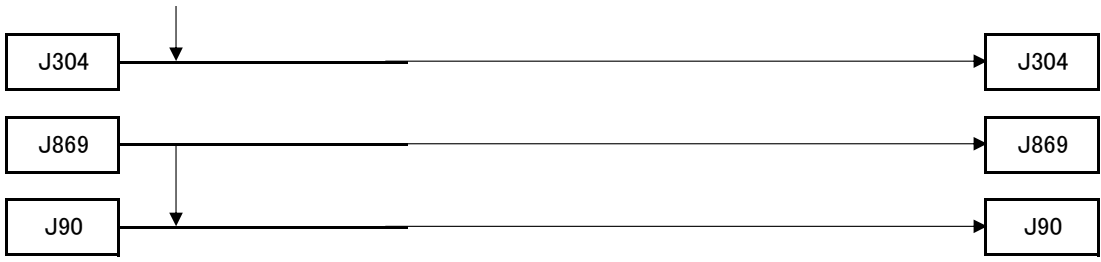


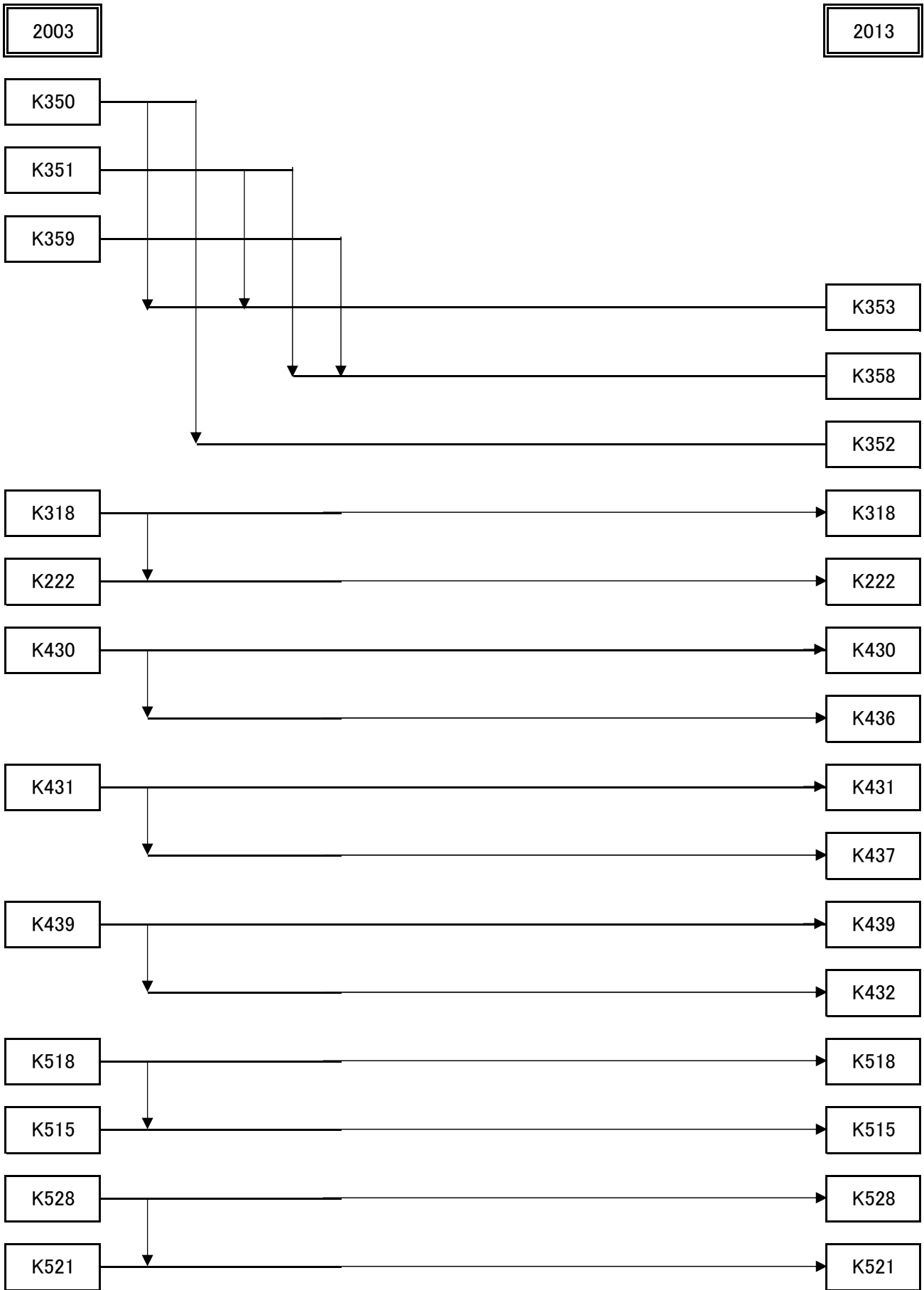


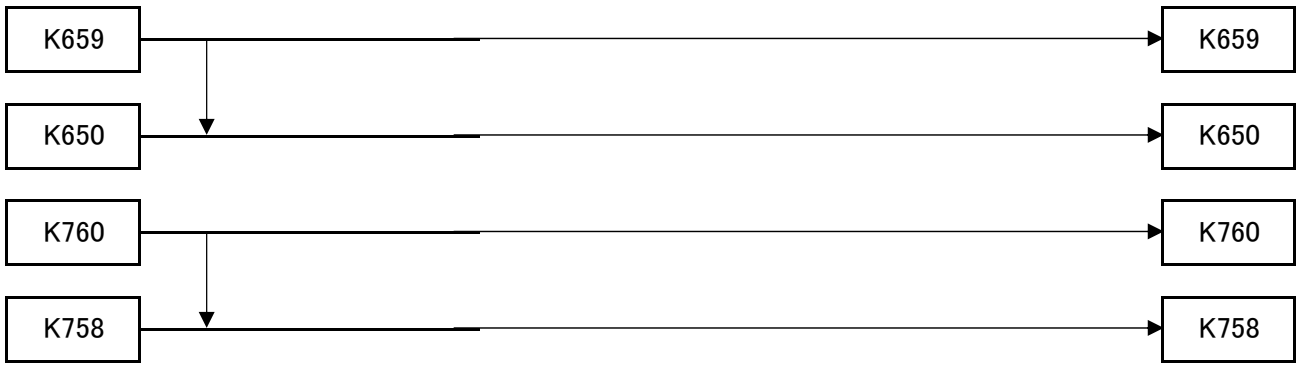
2003

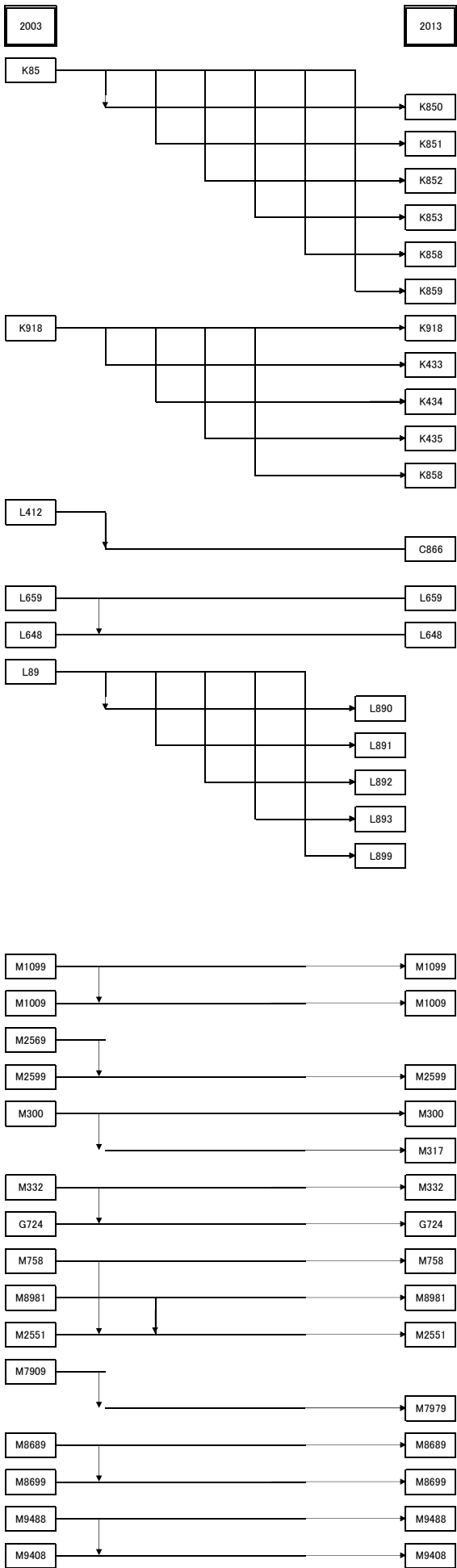
2013

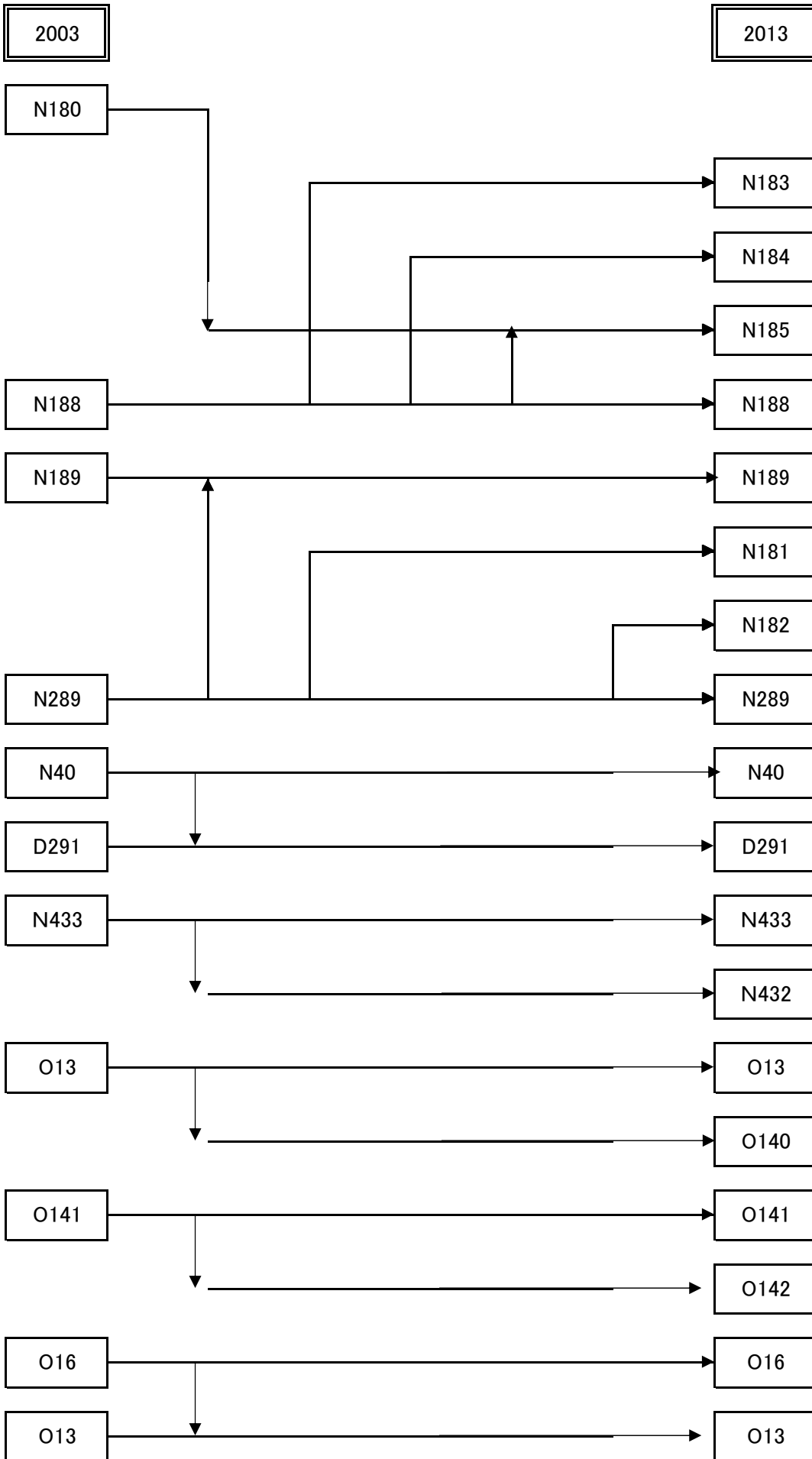


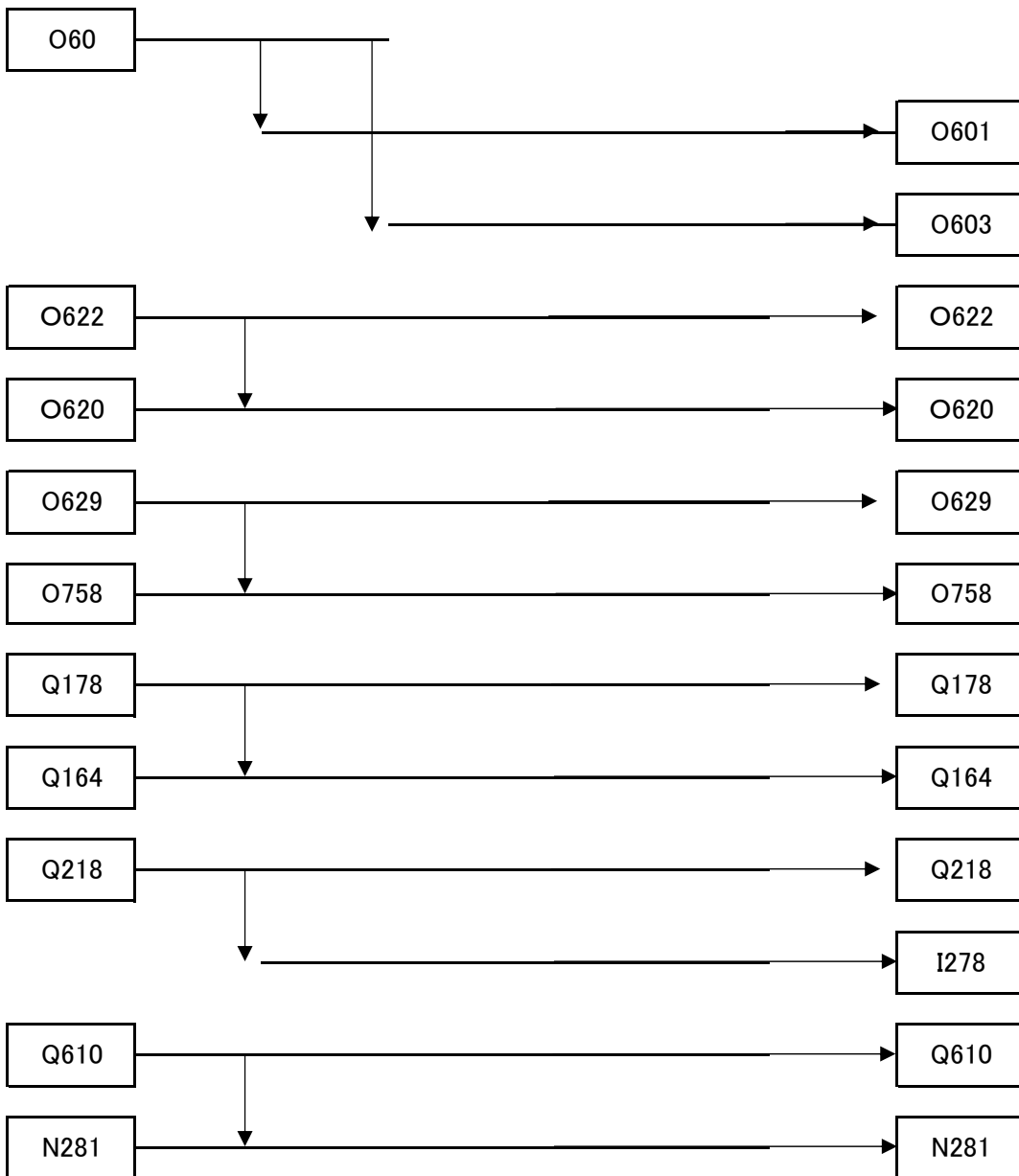


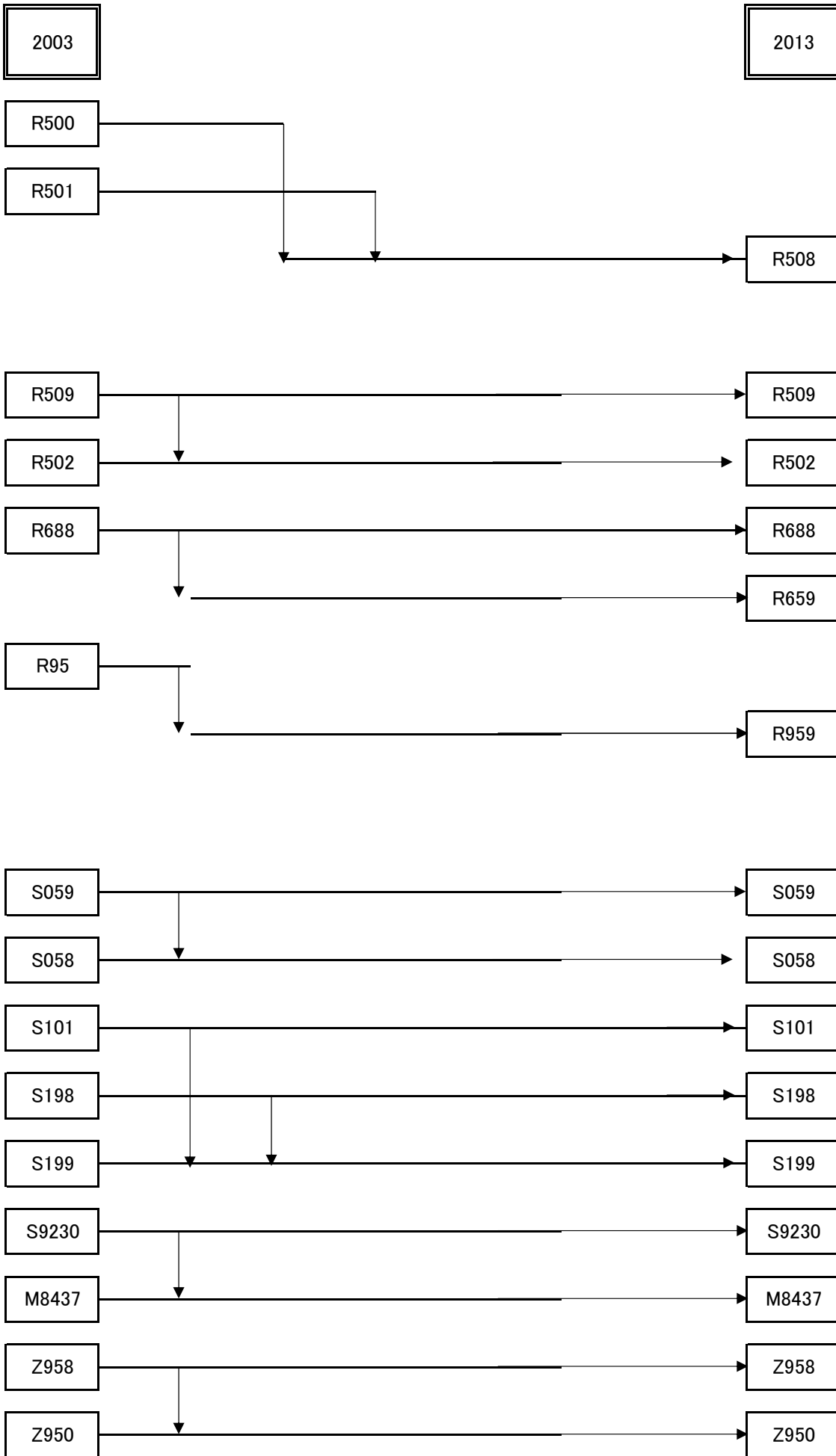


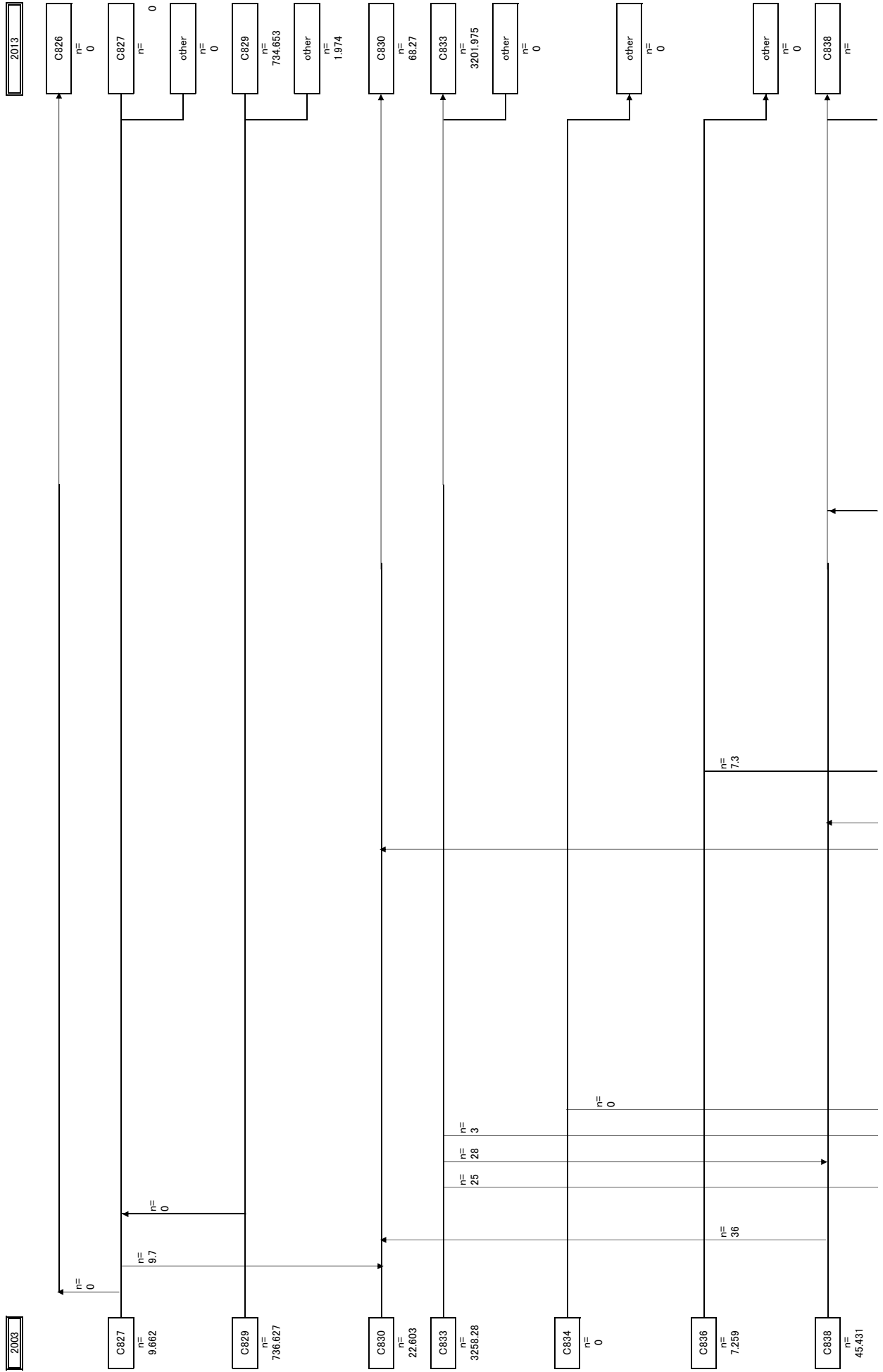


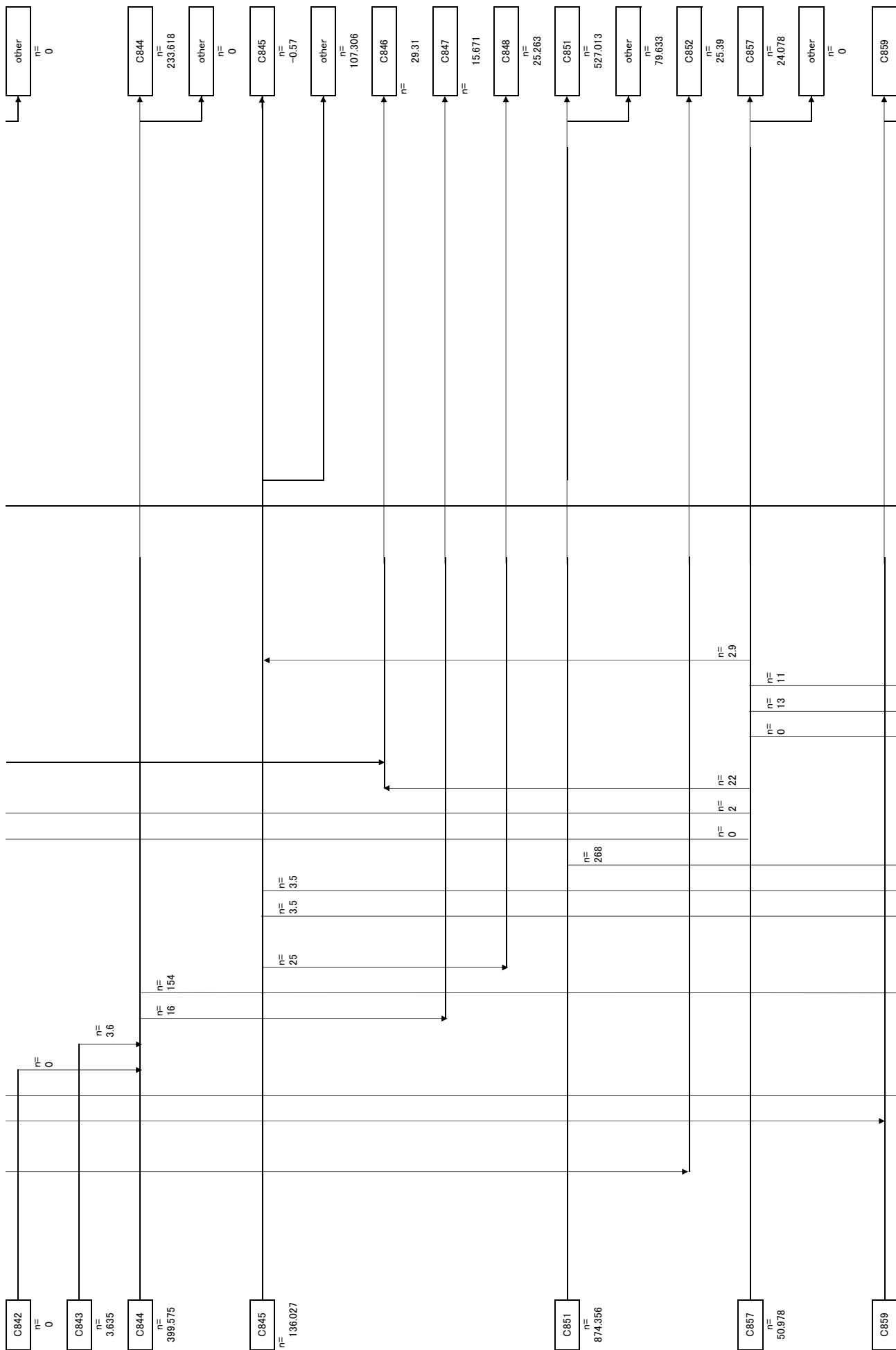


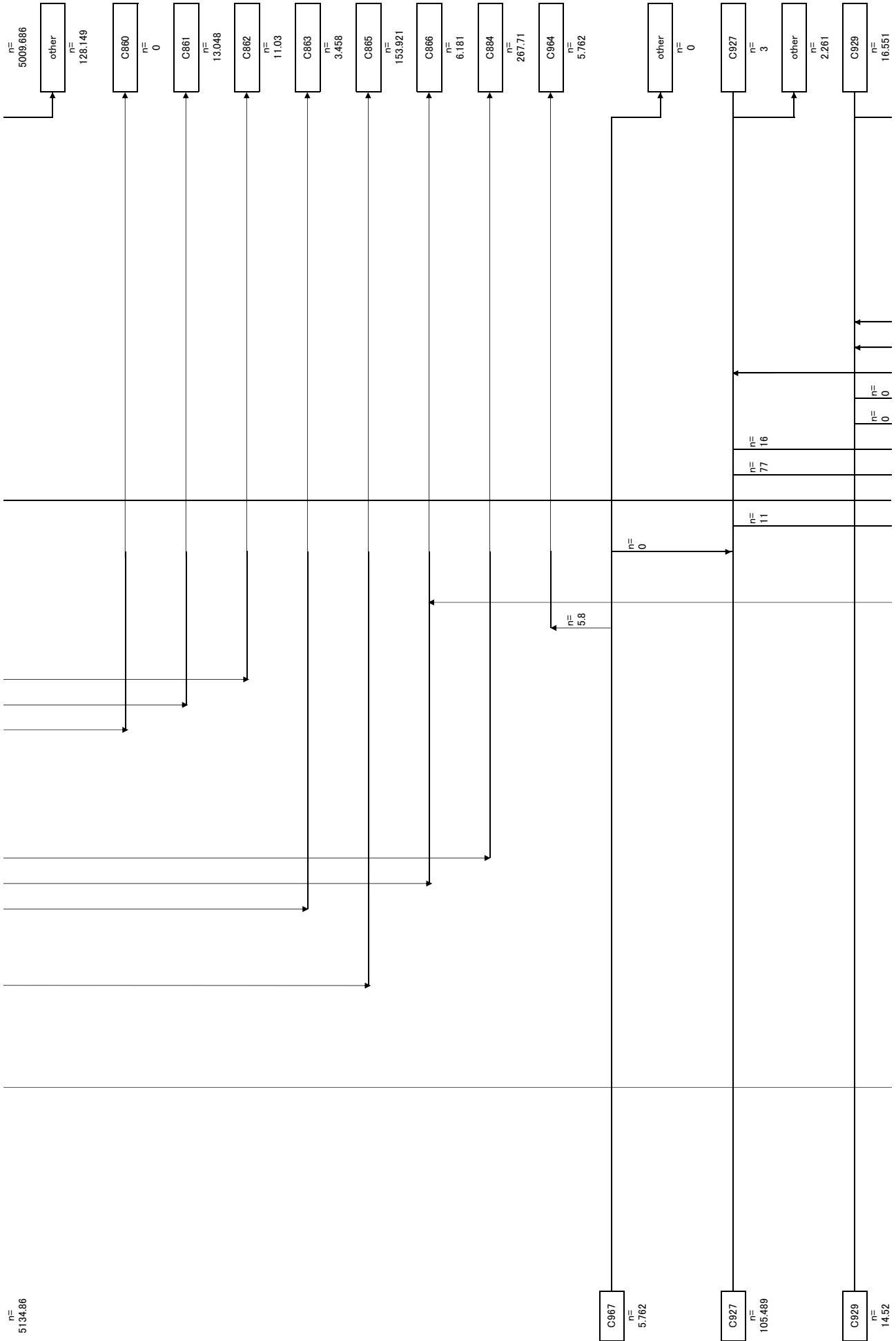


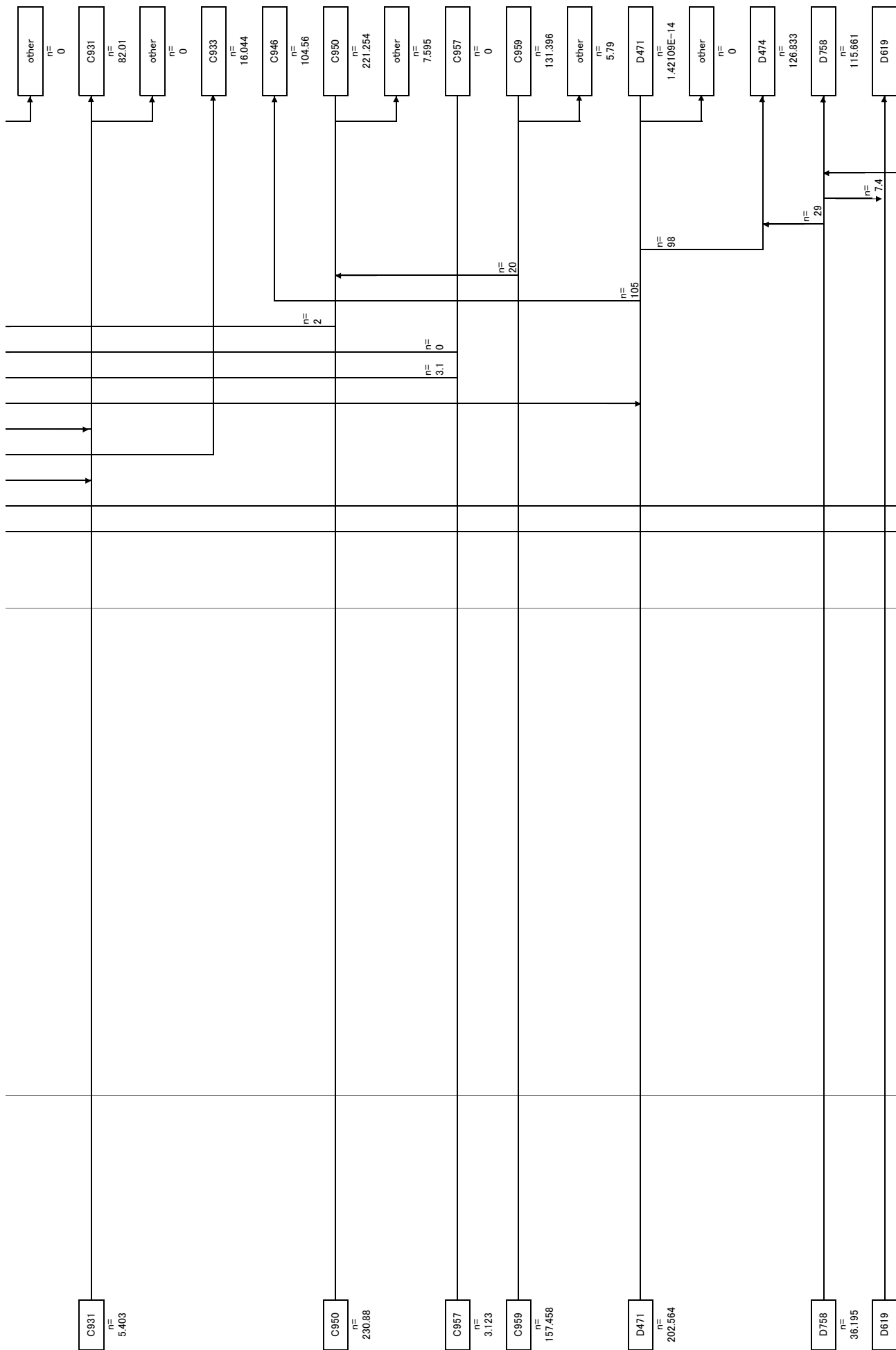


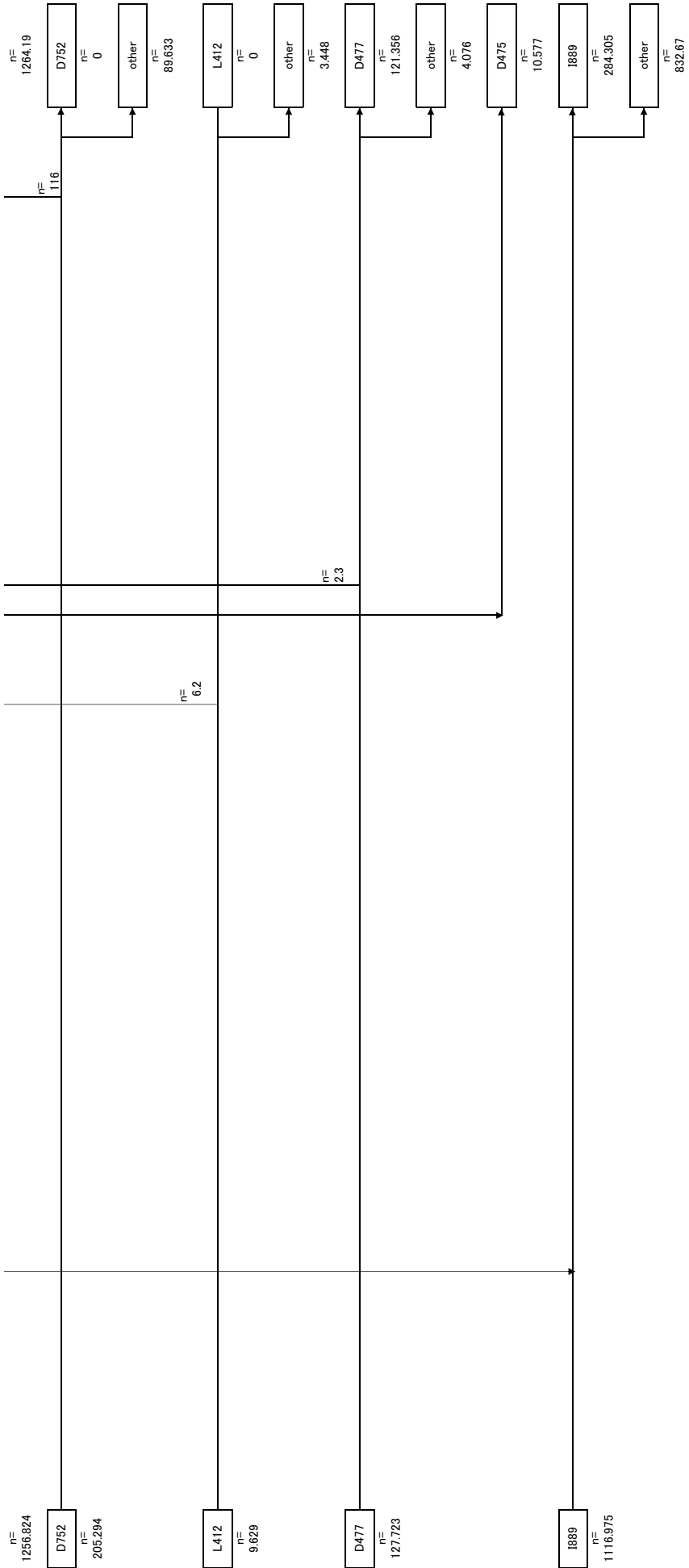


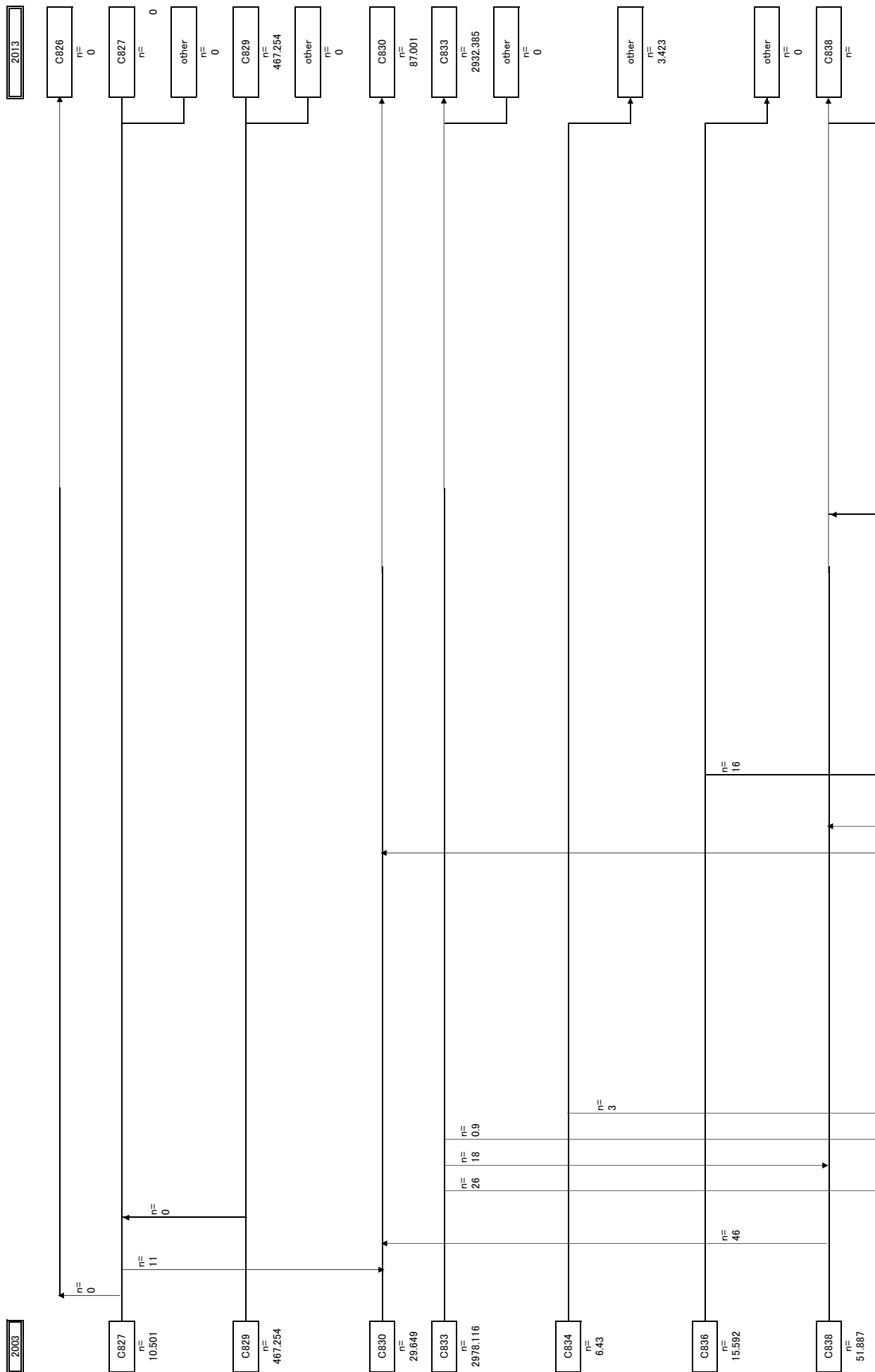


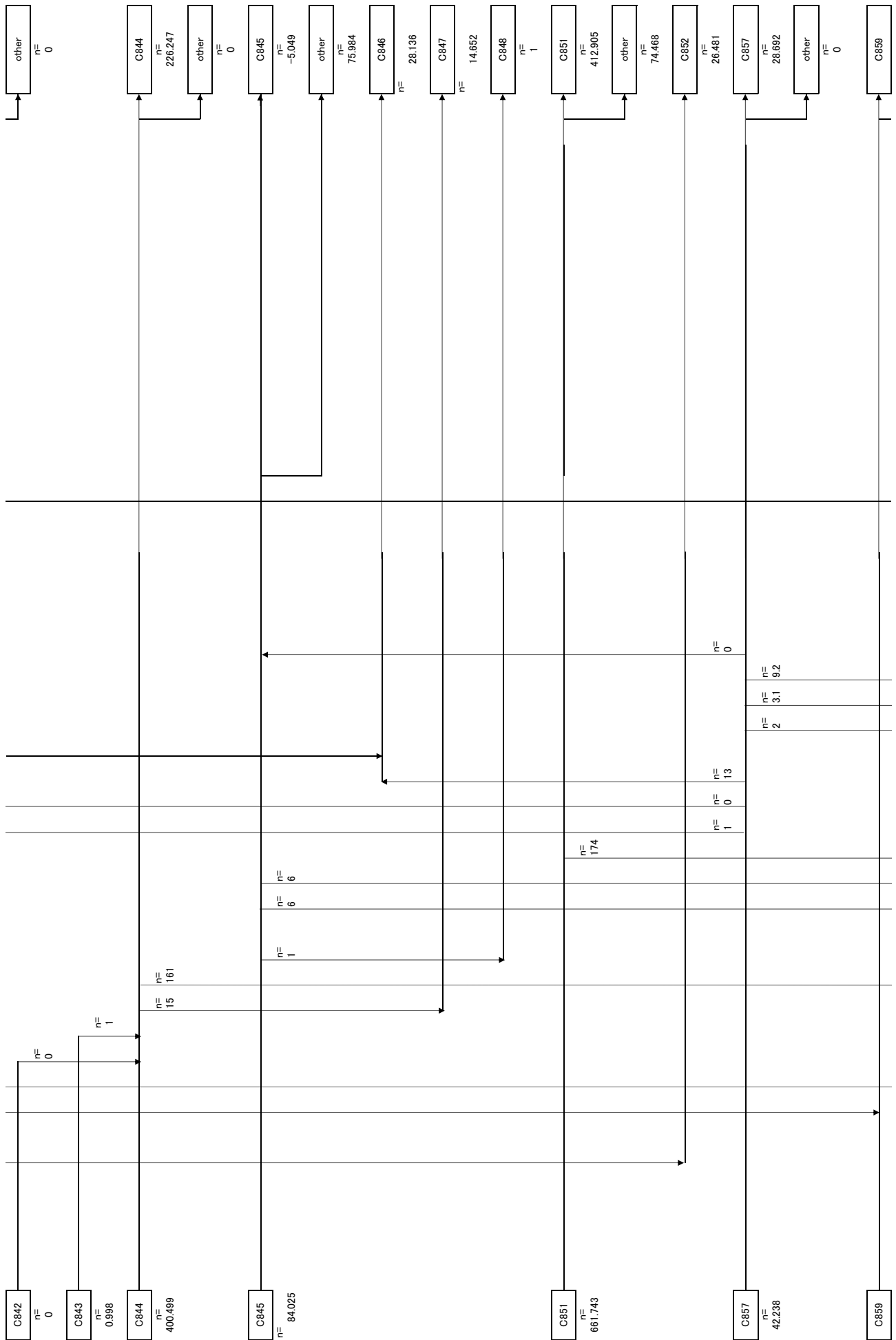


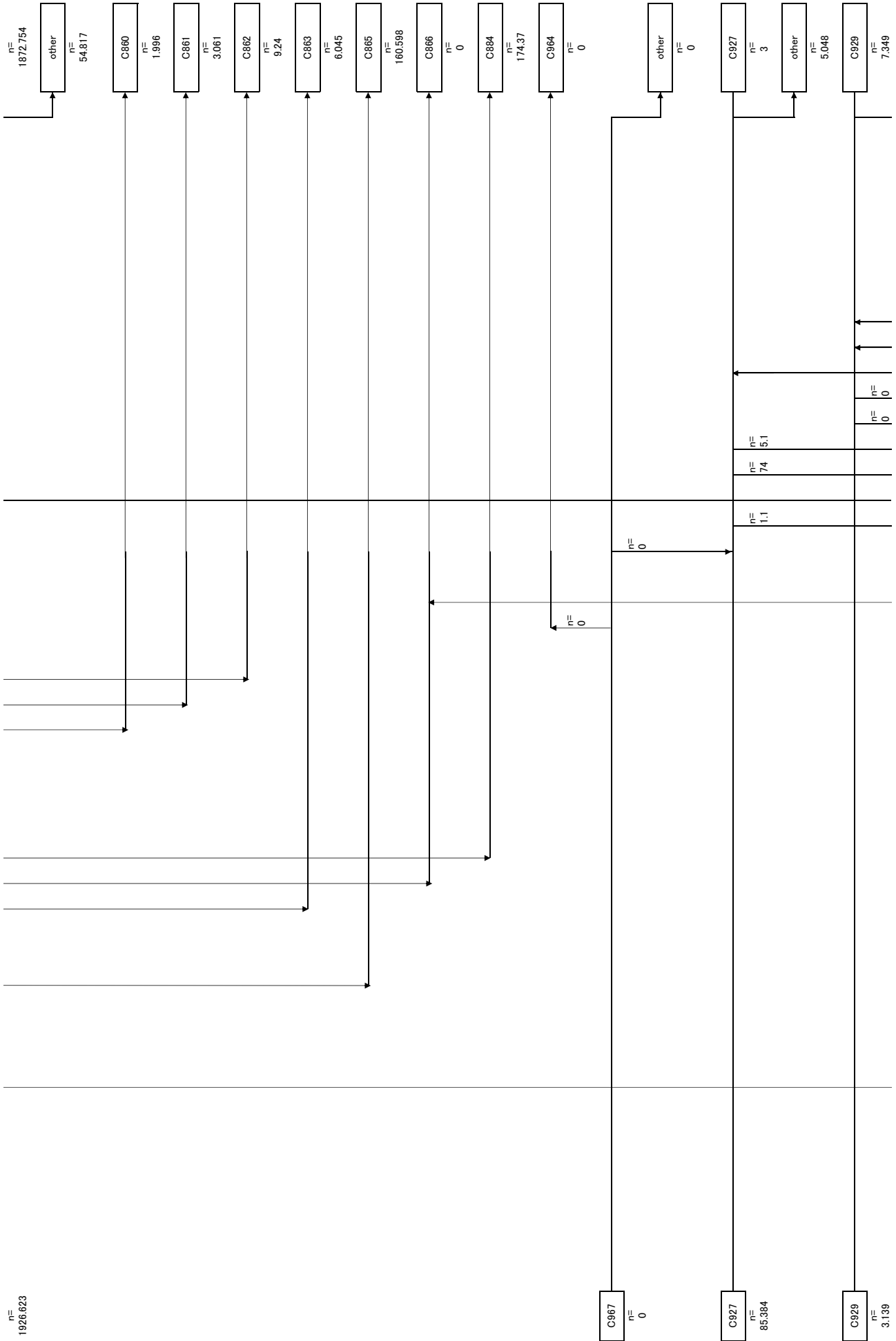




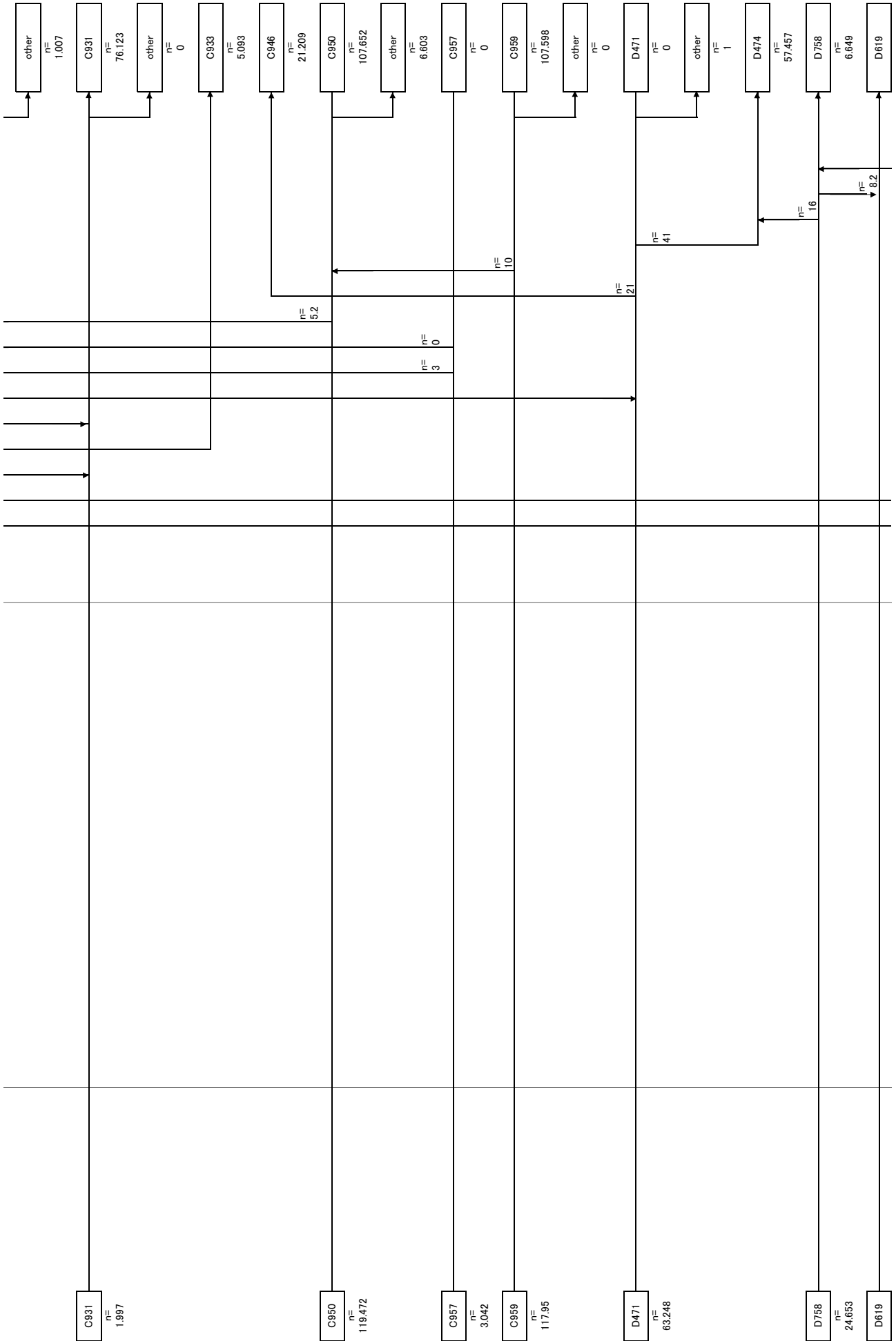


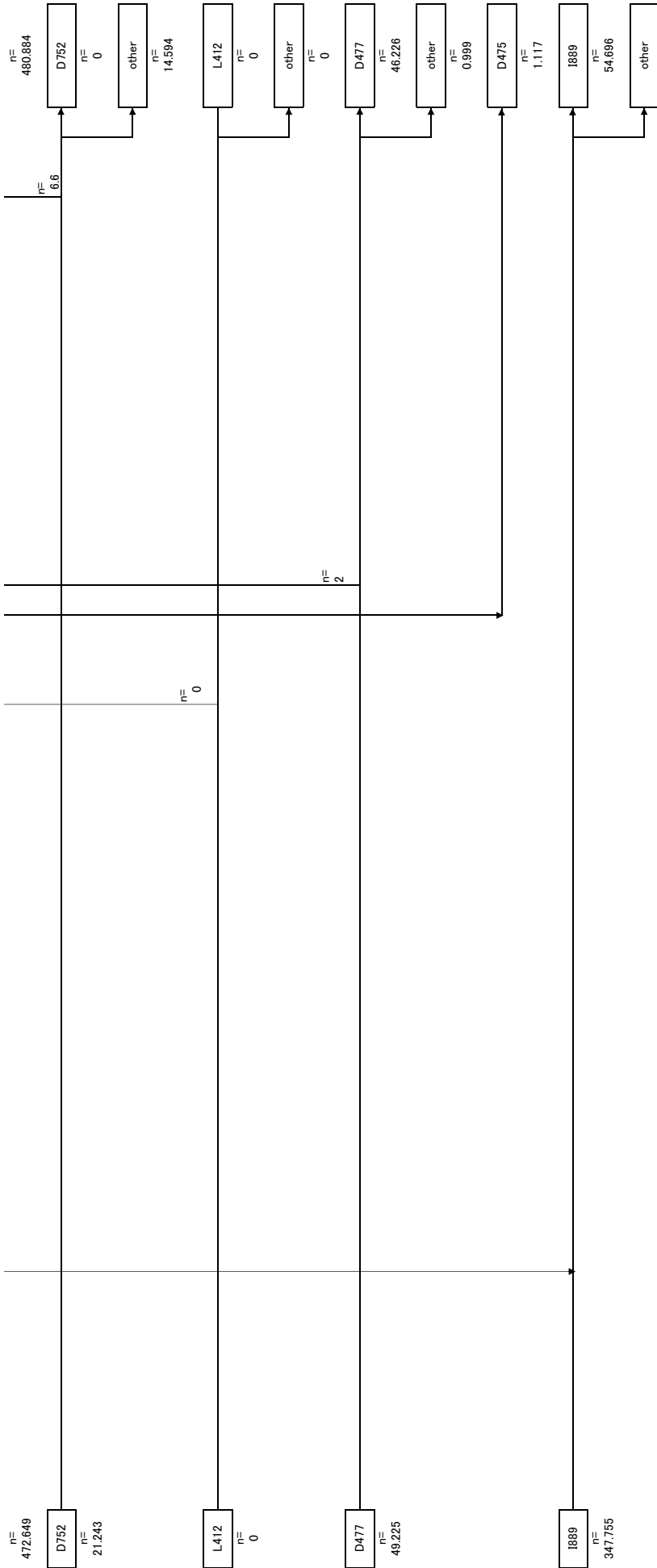




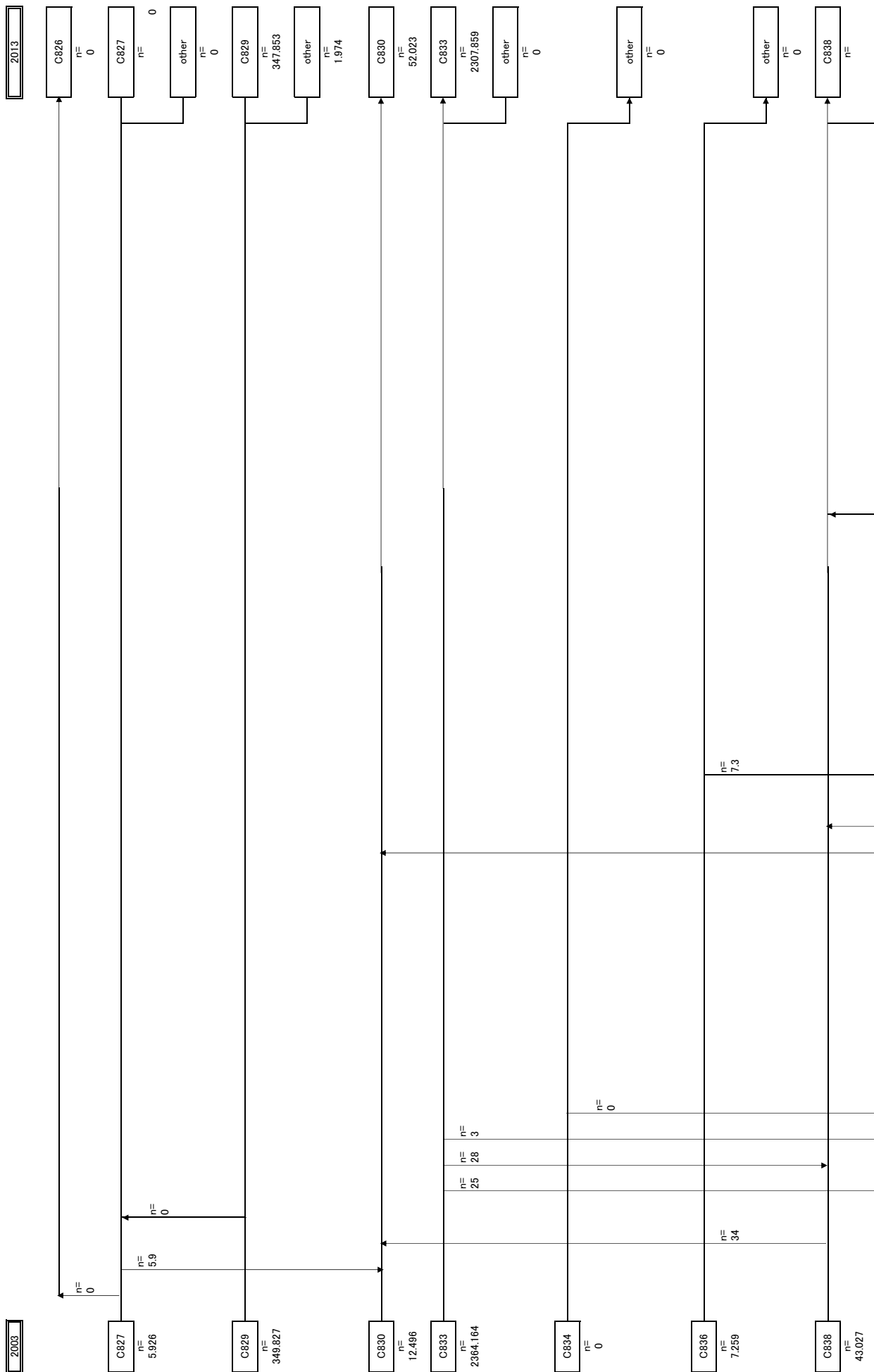


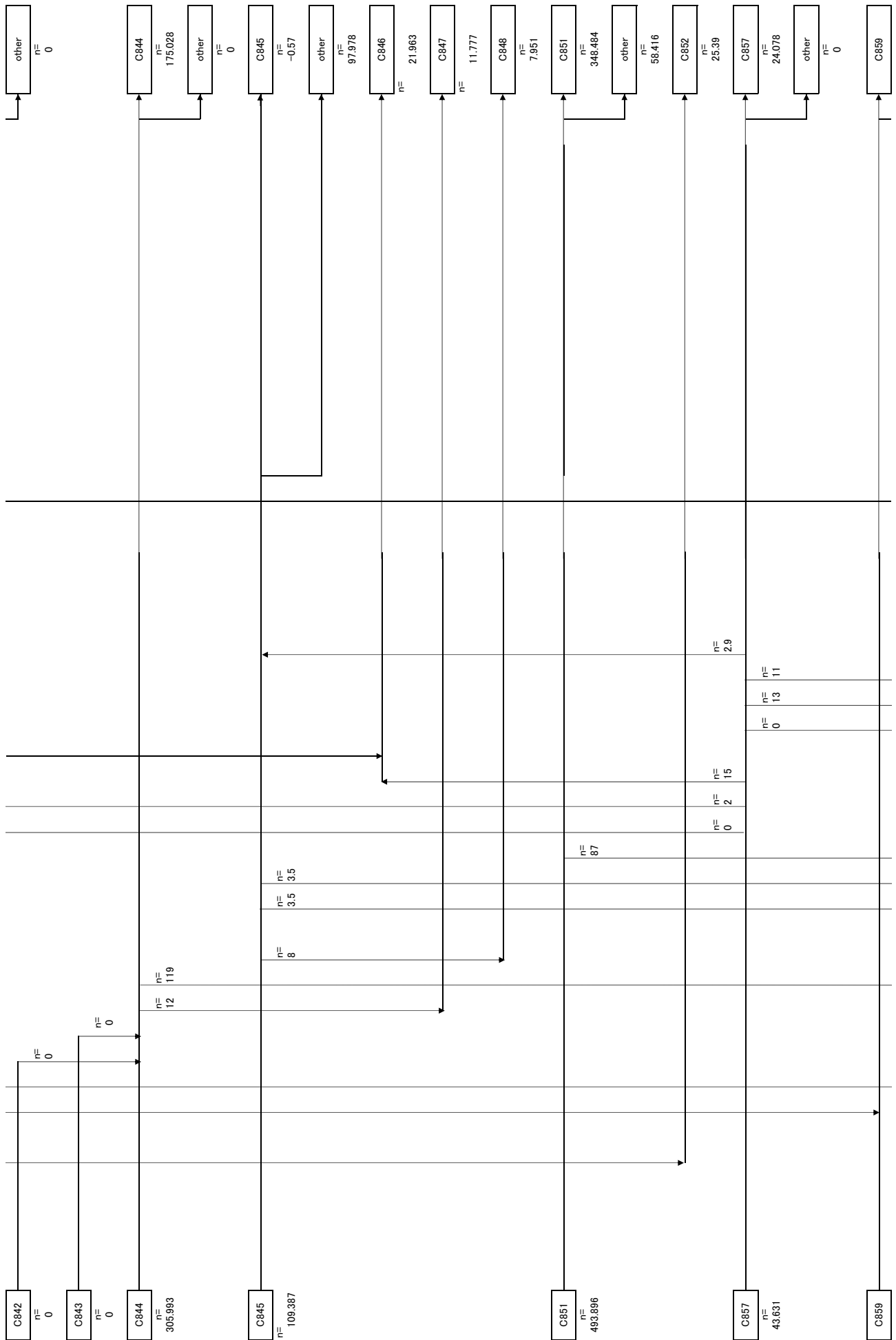
n= 1926.623

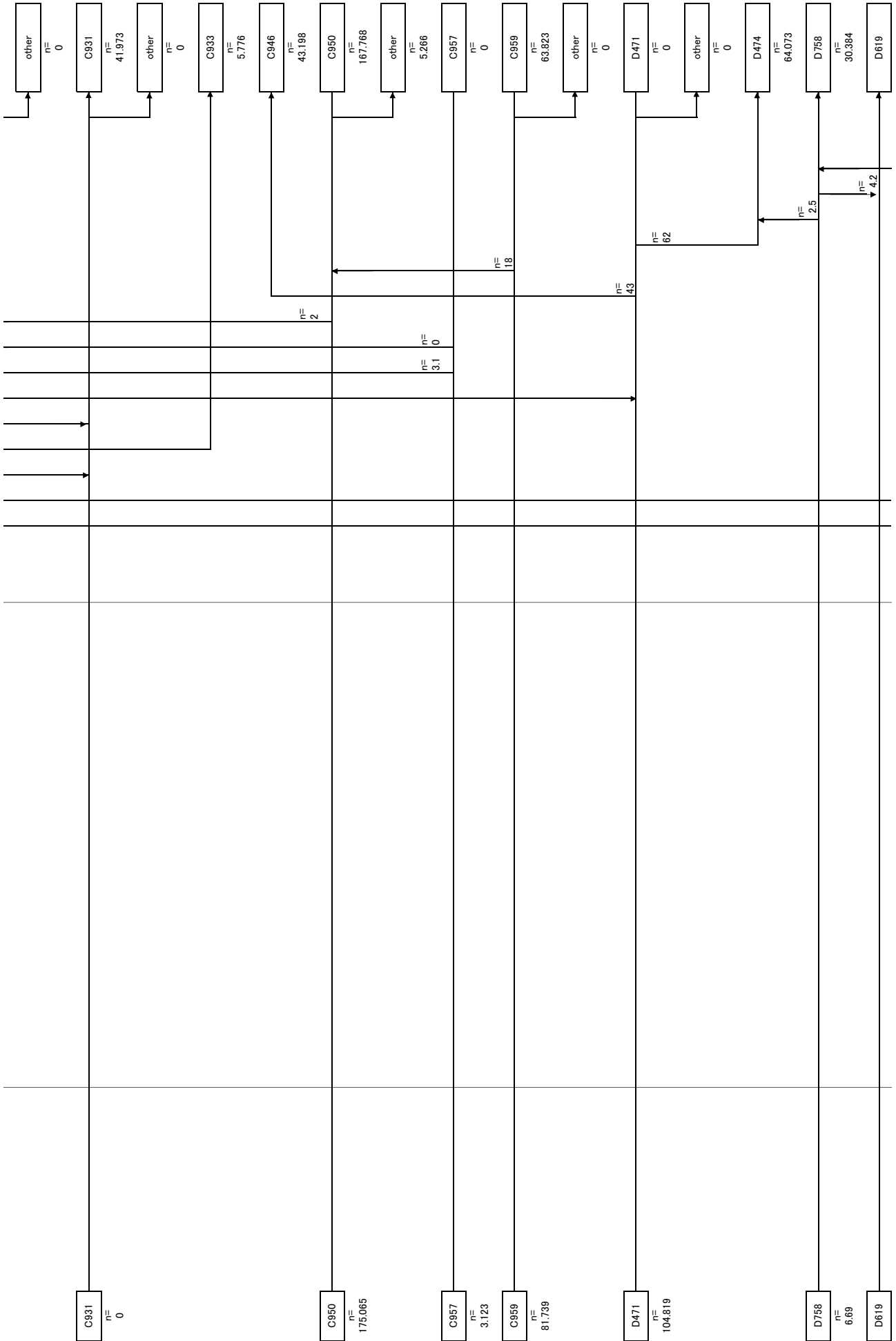


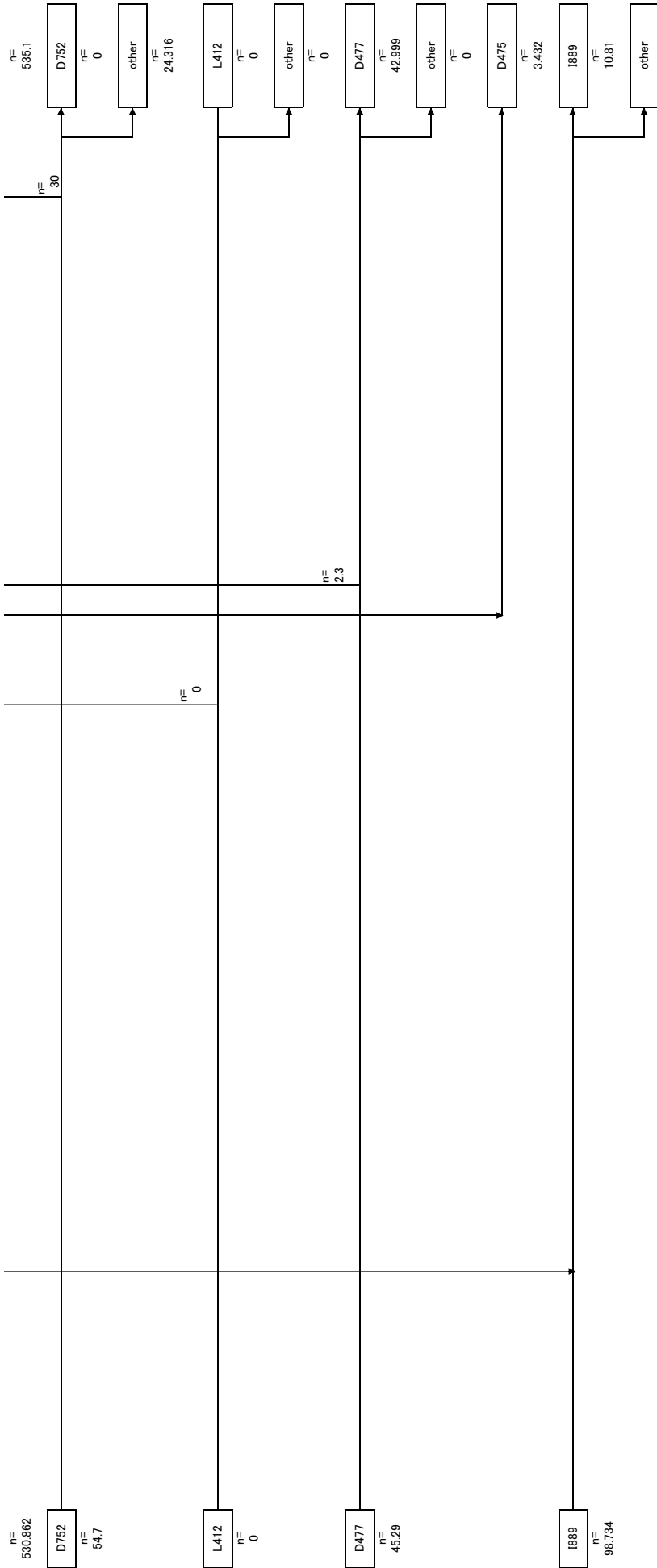


入院

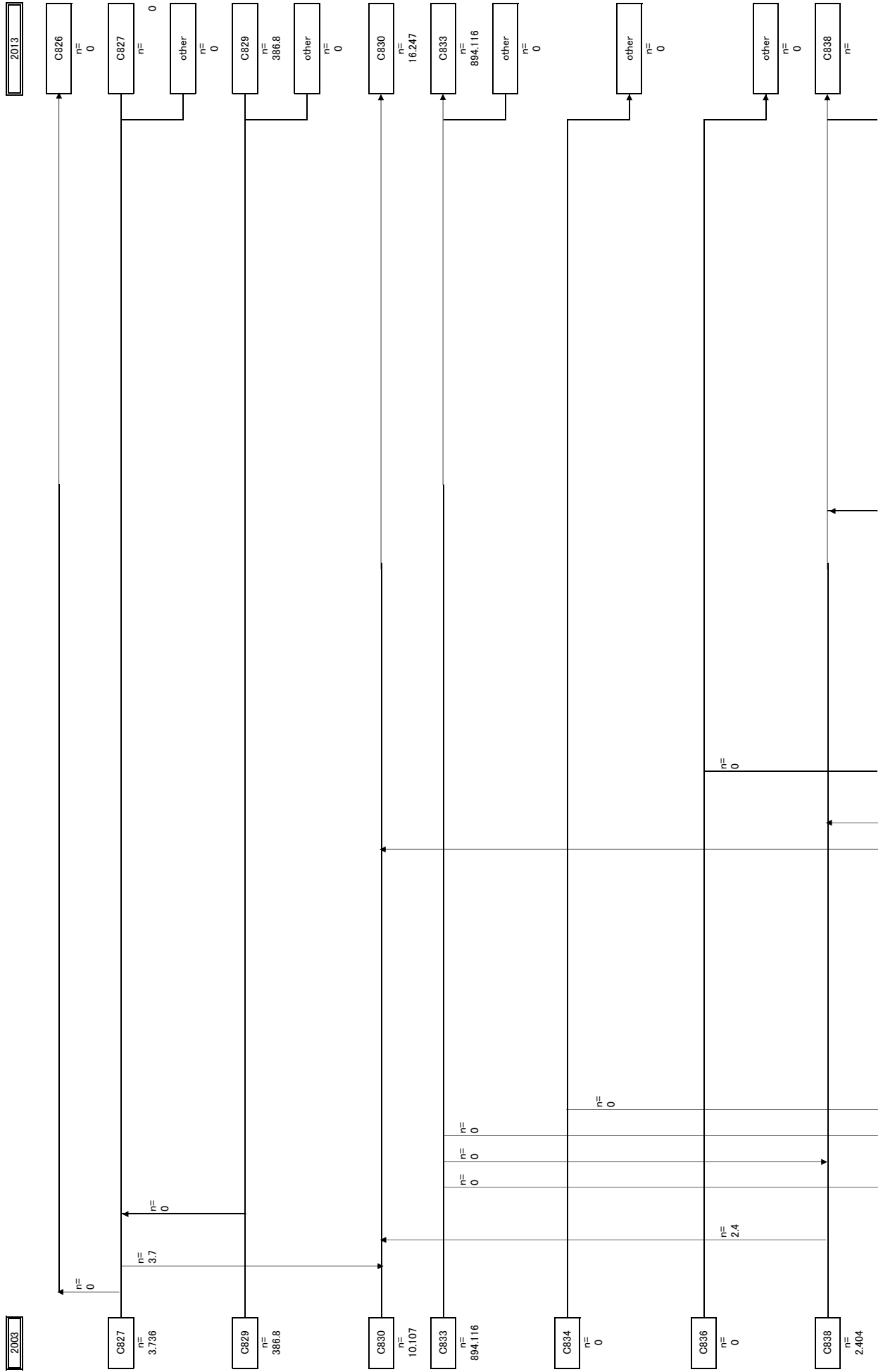


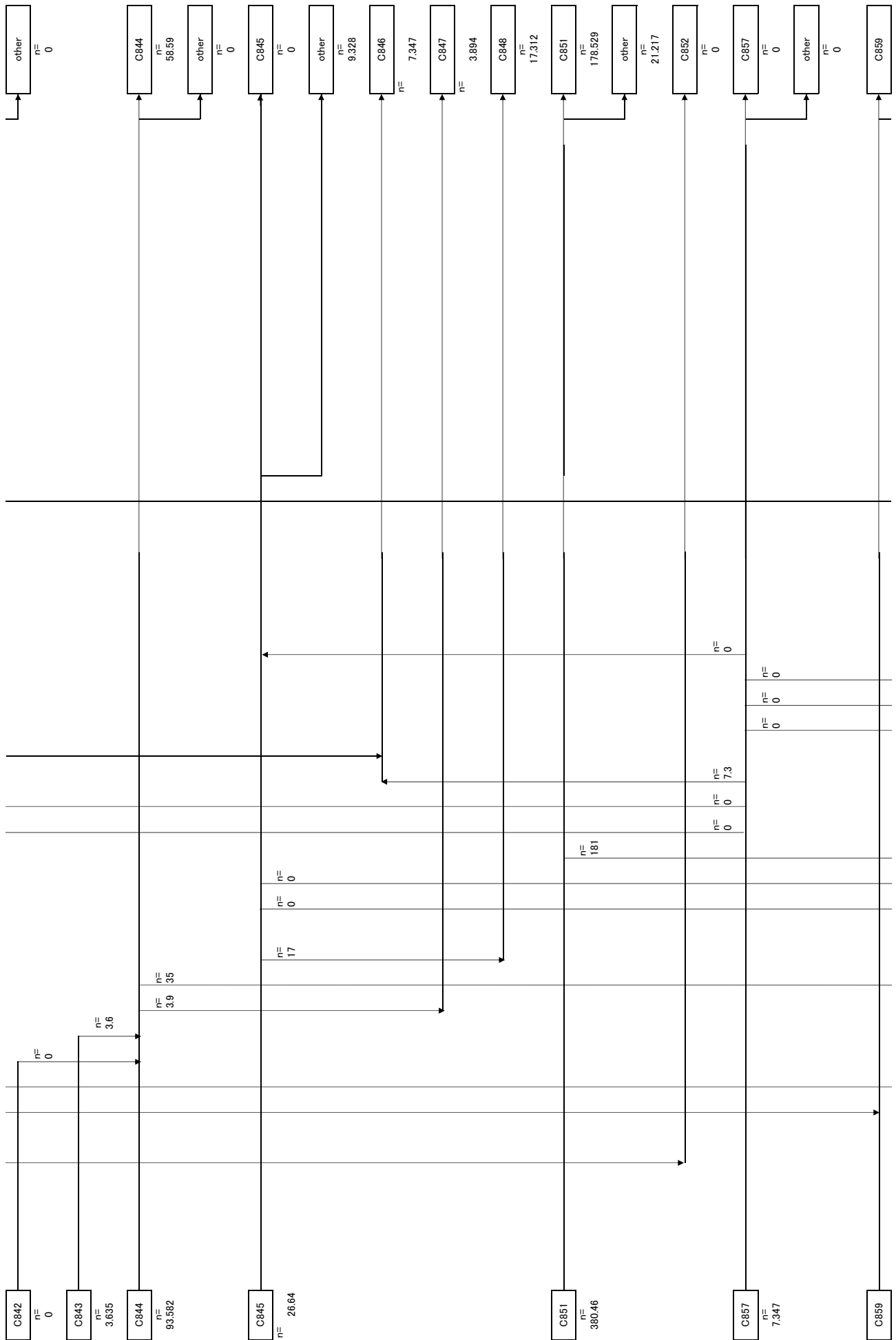




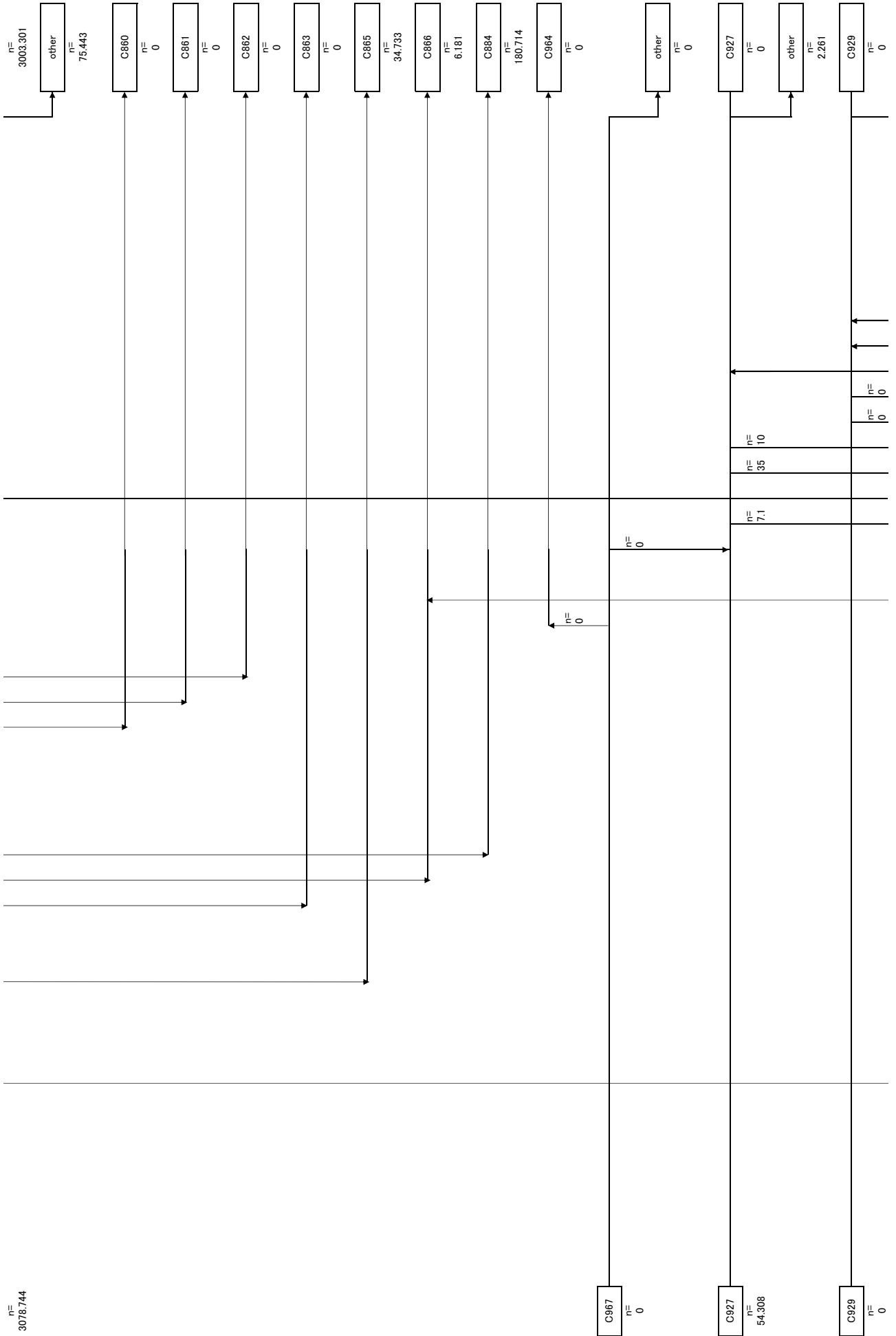


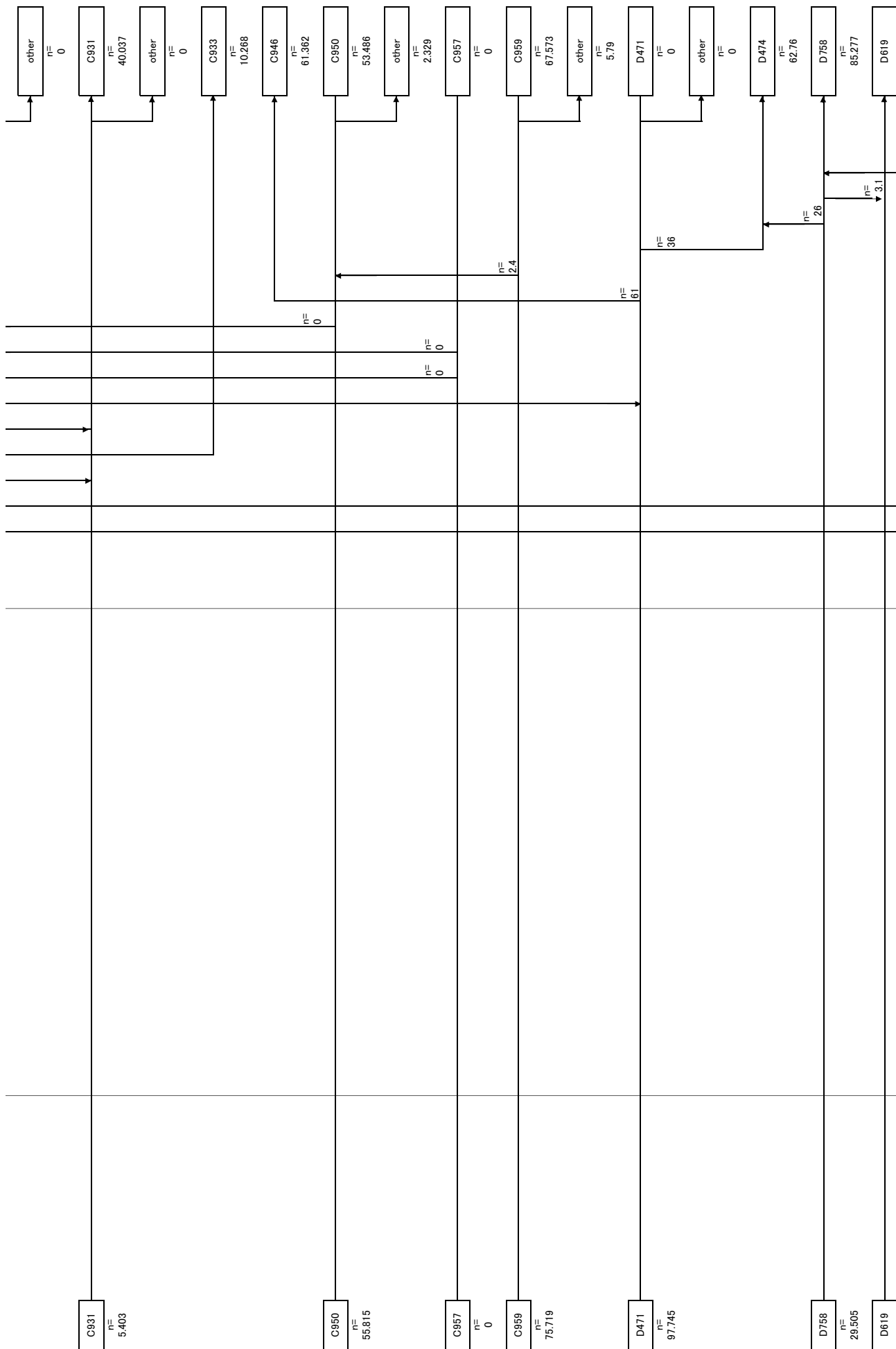
外来

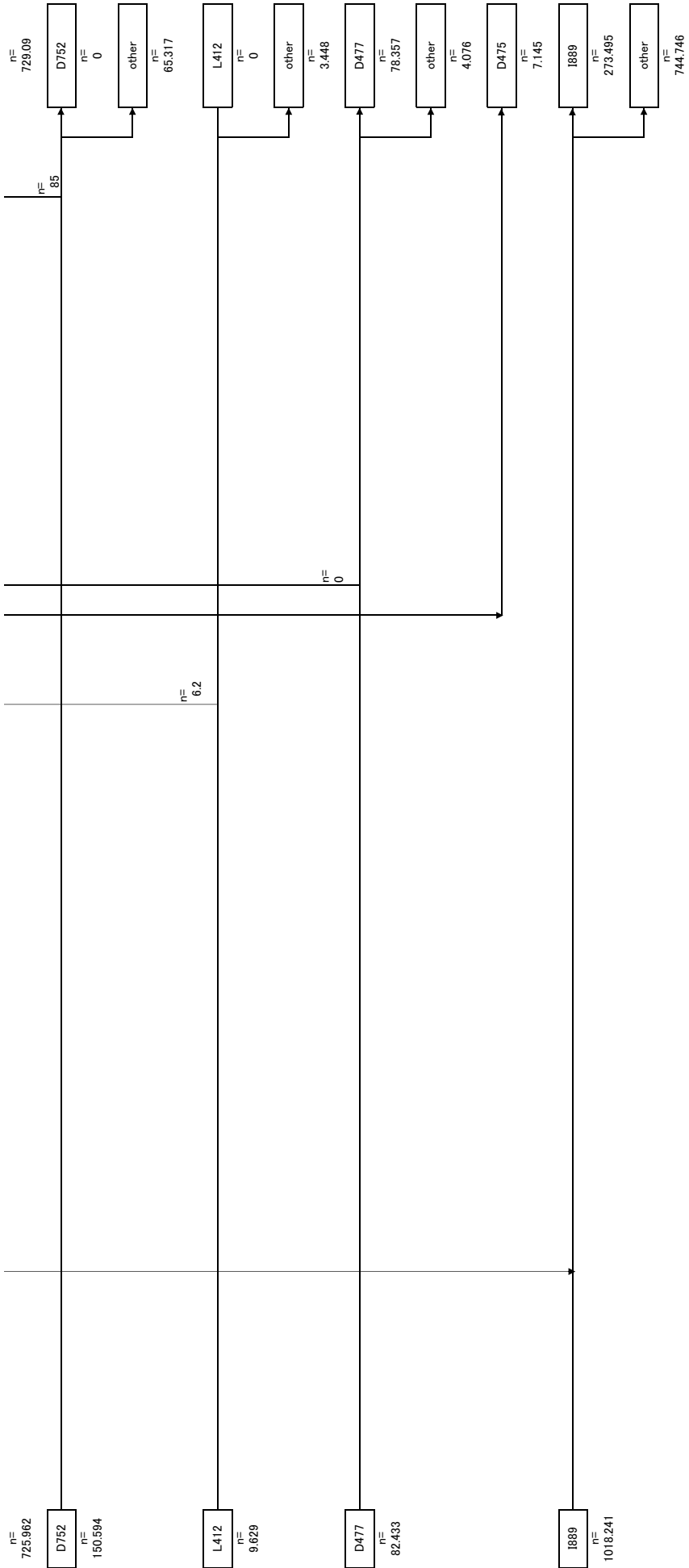




n= 3078.744







患者動態調査分類変更

データ解析手順

1. 死亡票

死亡票のデータは「市区町村符号および保健所符号」「事件簿番号」「男女別」「生年月日」「死亡したとき」「原死因符号」からなる。

「市区町村符号および保健所符号」「事件簿番号」の値を結合した変数を作成。

2. 死亡個票

死亡個票のデータは「市区町村符号および保健所符号」「事件簿番号」「死亡の原因（個票によって数はバラバラ）」「死因の種類」からなる。

各ファイルを csv に変換し、R で結合後、「市区町村符号および保健所符号」「事件簿番号」の値を結合した変数を作成。

3. 死亡票と死亡個票のマッチング

「市区町村符号および保健所符号」「事件簿番号」の値を結合した変数を用いて、死亡票と死亡個票をマッチングさせる。

マッチング率は 93.00%だった。

4. 分類変更コードのマッチング

3 のマッチングデータのうち、分類変更コードデータを抽出（コードでマッチング）。

以下は各コードの分類変更集計

5. 「C859」の分類変更について集計する（マッチング後 9056 個）

・ C859 の変更に関与するコードは

C830、C833、C836、C838、C845、C857、C827、C834、C836、C842、C843、C844、C851、C967、C927、L412、D477、I889、C927、C929、C931、C957、C959、D471、D619、D752、D758

・ D コードは今回のデータに個票なし

・ data0.csv より上記のコードを抽出し、死亡原因を抜き出す

・ 病名完全一致の条件で分類変更集計したパターンと、目視による調整を行い分類編子集計したパターンを示す

6. 「C80」の分類変更について集計（マッチング後 4397）

・ C80 の変更に関与するコードは C97

・ 上記のコードを抽出し、死亡原因を抜き出す

- ・病名完全一致の条件で分類変更集計したパターンと、目視による調整を行い分類編子集計したパターンを示す

- ・C80 の再コーディングは非常に困難。明確なコーディング基準を決めることができない、死亡原因 1 - 4 の優先順位の判断（死亡原因 1 を優先するのか、正しい病名記載の死亡原因を優先するのか、など）。コーディングエラーの影響が強く出ると考えられる。

- ・C97 は「重複癌」であるが、複数の癌を列挙しているものも「重複癌」としてよいのかどうか、など判断が困難。

- ・再コーディングが困難なため、病名が完全一致する方法で分類変更集計を試みたが、そうすると C97 はほとんどが「other(分類不可)」になってしまう。

7. 「C920」 の分類変更について集計

- ・ C920 の変更に関与するコードは D463

- ・ D463 は今回のデータに含まれない

8. 「C900」 の分類変更について集計

- ・ C900 の変更に関与するコードは C902

- ・ 上記のコードを抽出し、死亡原因を抜き出す

- ・病名完全一致の条件で分類変更集計したパターンと、目視による調整を行い分類編子集計したパターンを示す

9. 「C915」 の分類変更について集計

- ・「ICD 分類パターン解析_人口動態_C915_C947.xlsx」 参考

- ・ C915 の変更に関与するコードは C947

- ・ 上記のコードを抽出し、死亡原因を抜き出す

- ・病名完全一致の条件で分類変更集計したパターンと、目視による調整を行い分類編子集計したパターンを示す

10. 「C833」 の分類変更について集計

- ・ C833 は C859 の分類変更に関与するものですでに集計済み

11. 「C921」 の分類変更について集計

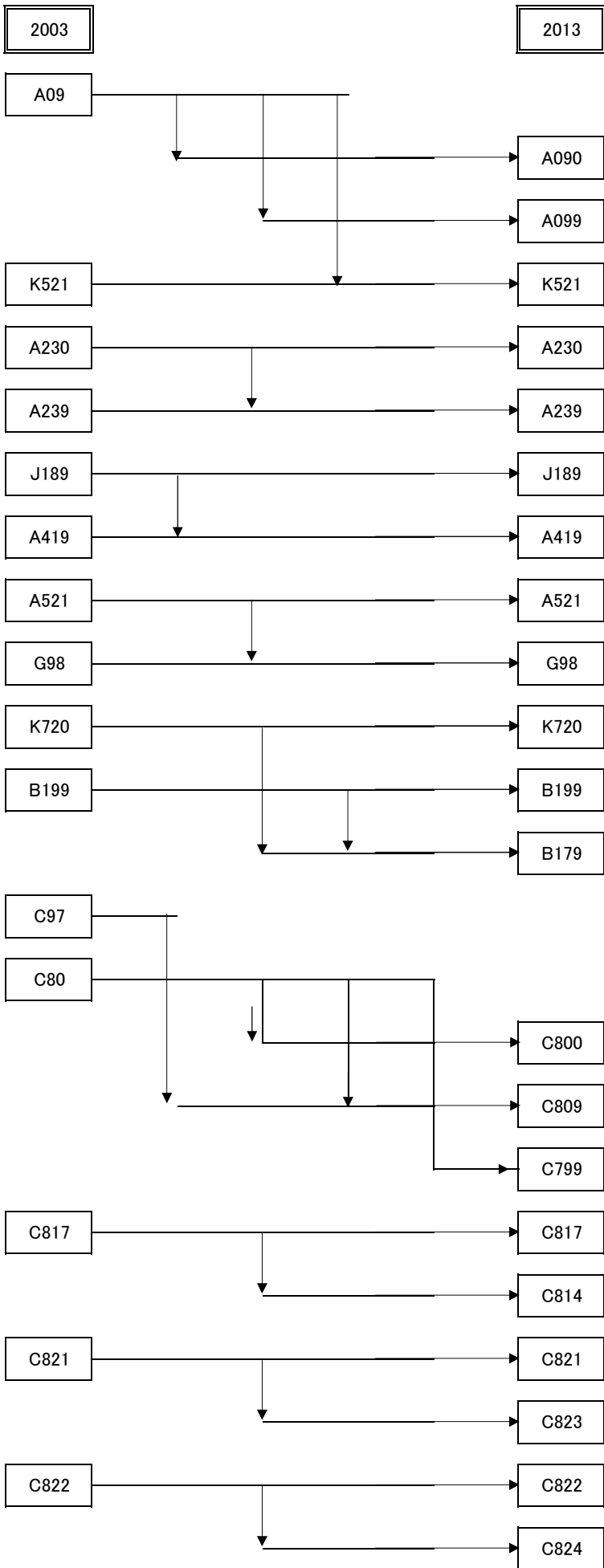
- ・ C921 の変更に関与するコードは C922

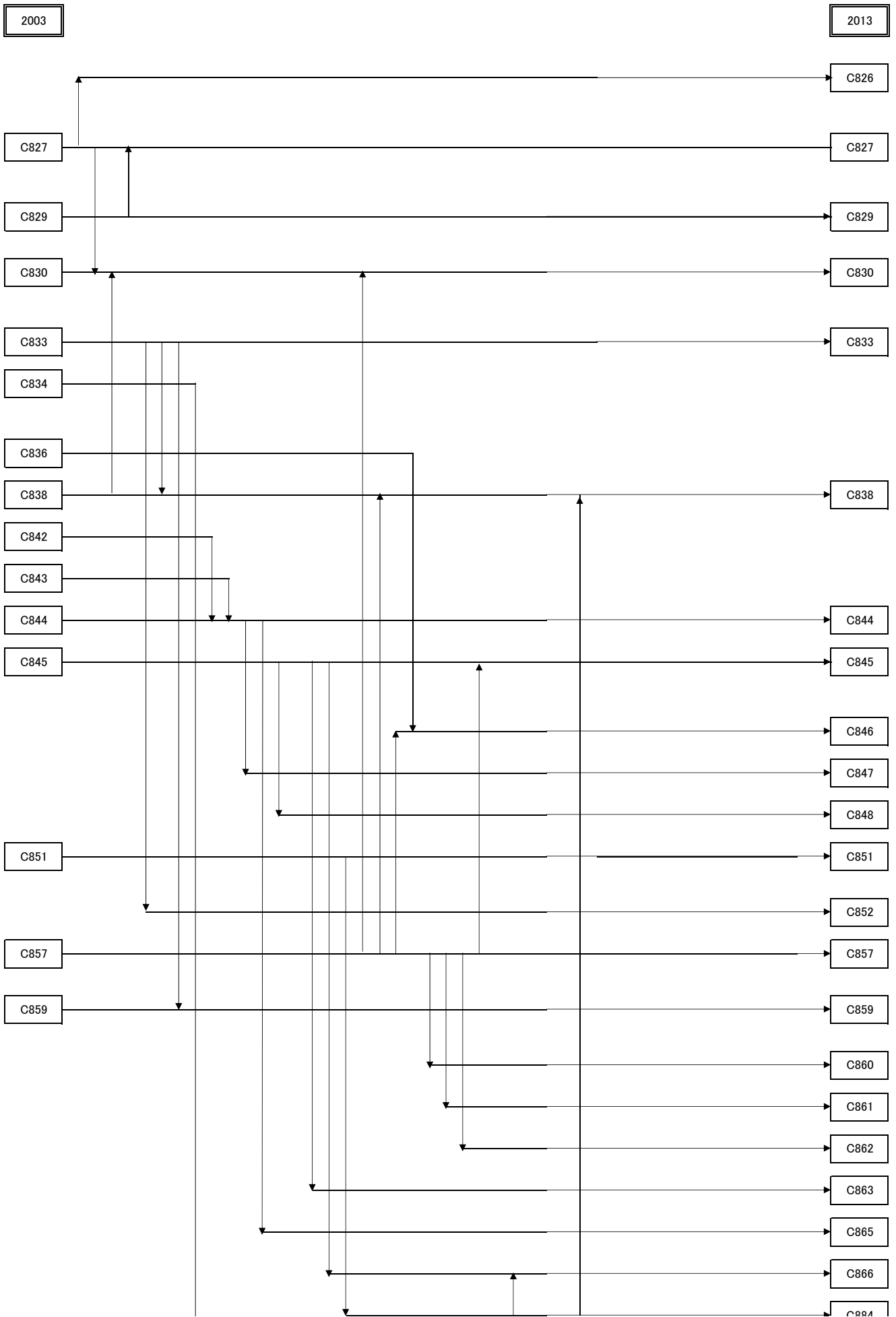
- ・ 上記のコードを抽出し、死亡原因を抜き出す

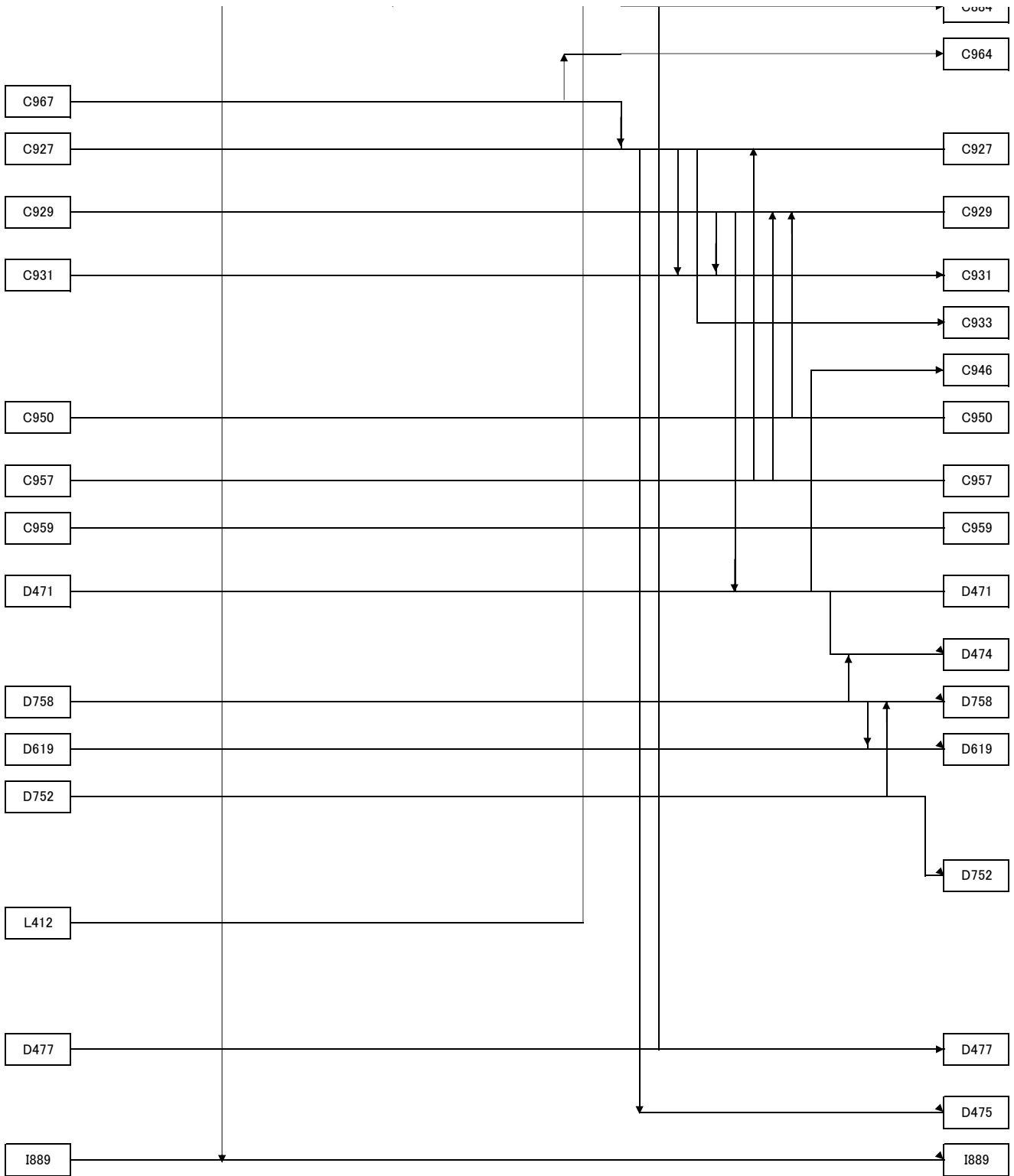
- ・病名完全一致の条件で分類変更集計したパターンと、目視による調整を行い分類編子集計したパターンを示す

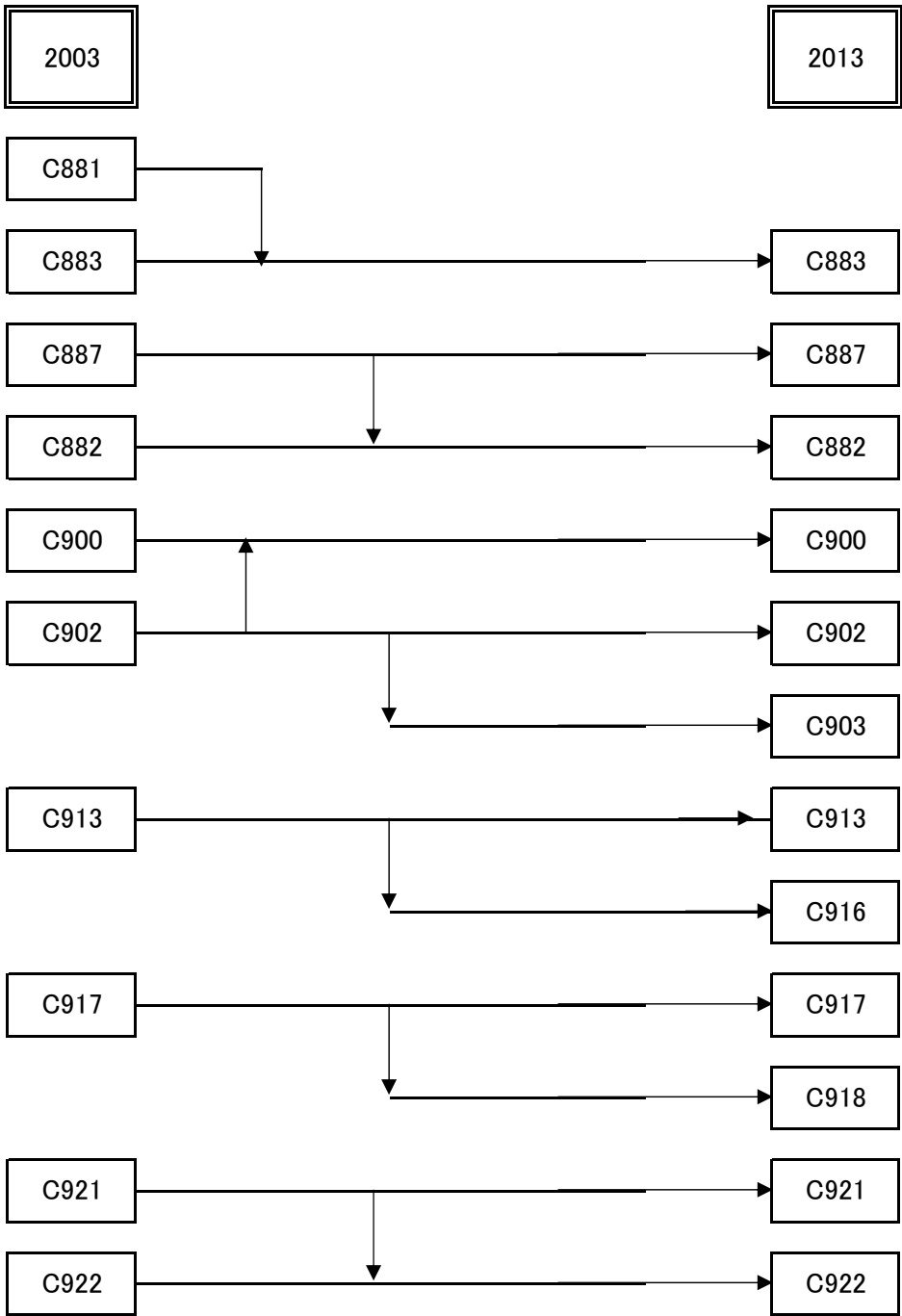
12. 「C961」の分類変更について集計

- ・上記のコードを抽出し、死亡原因を抜き出す
- ・病名完全一致の条件で分類変更集計したパターンと、目視による調整を行い分類編子集計したパターンを示す

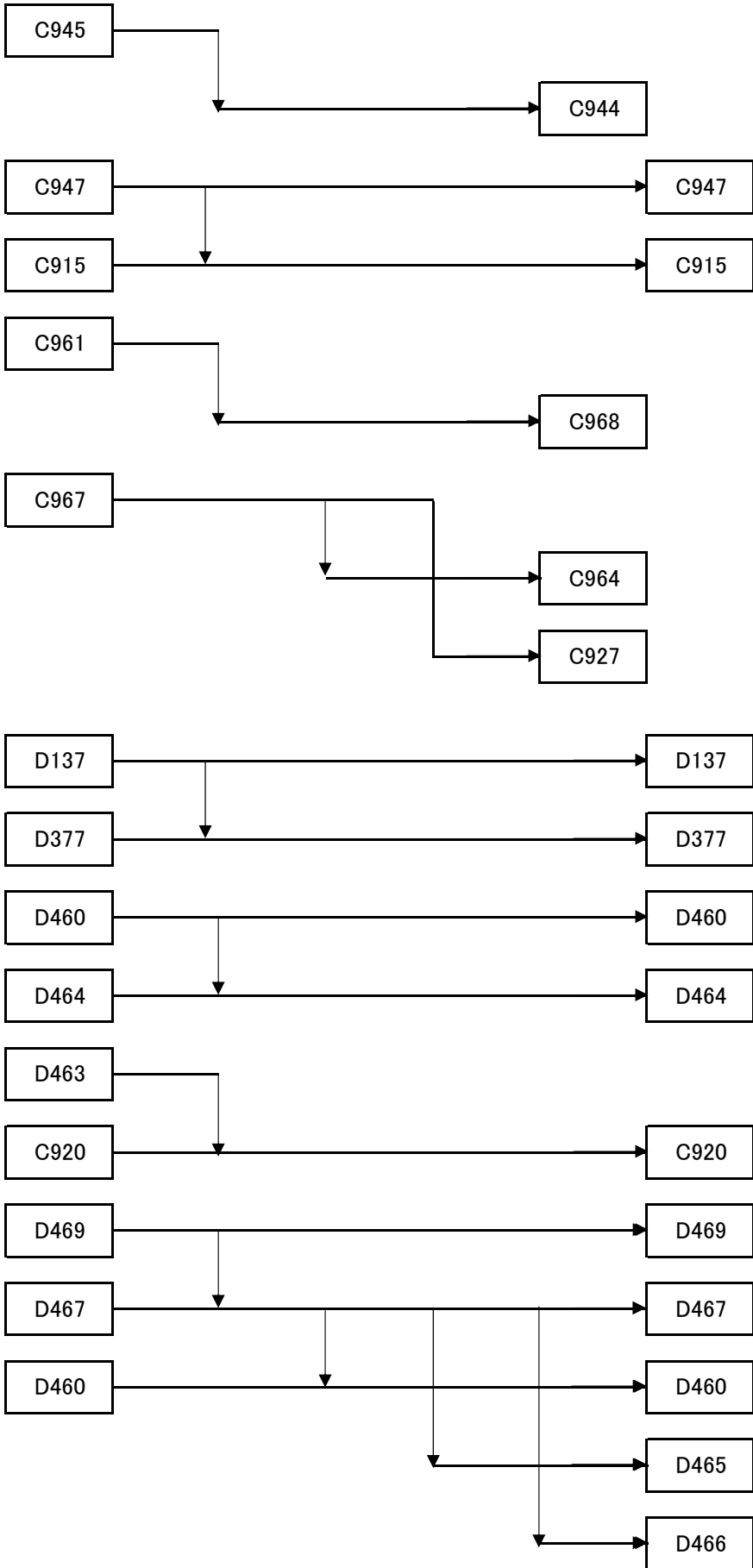


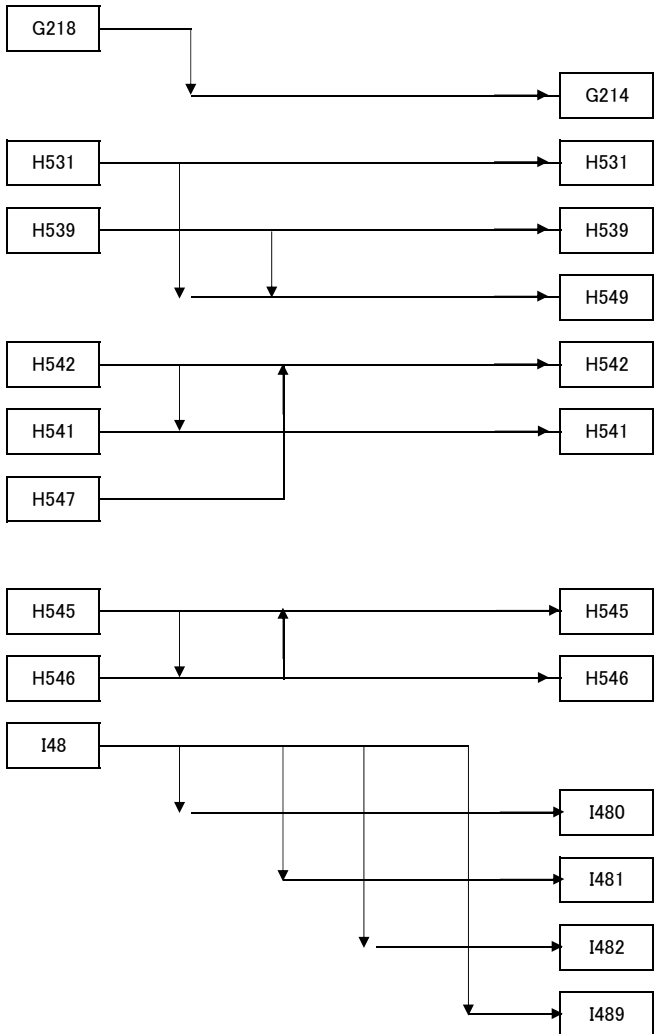
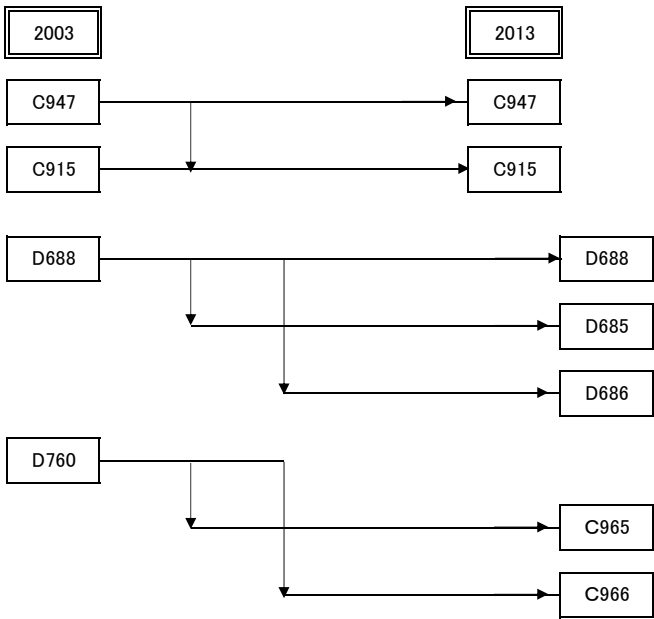


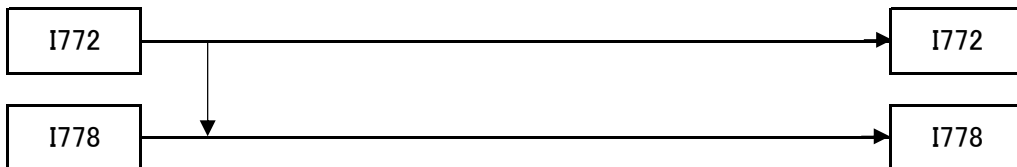
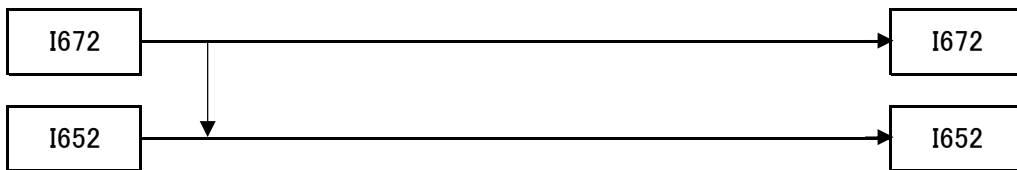
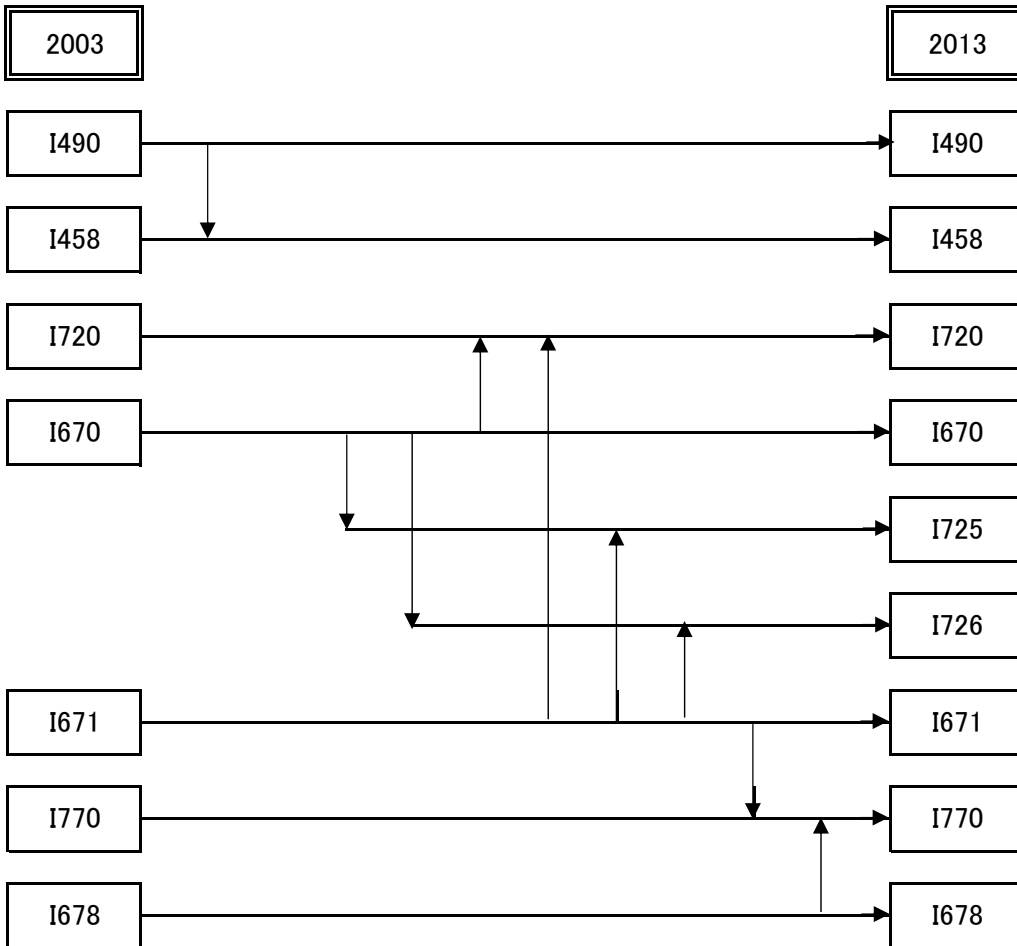




2003

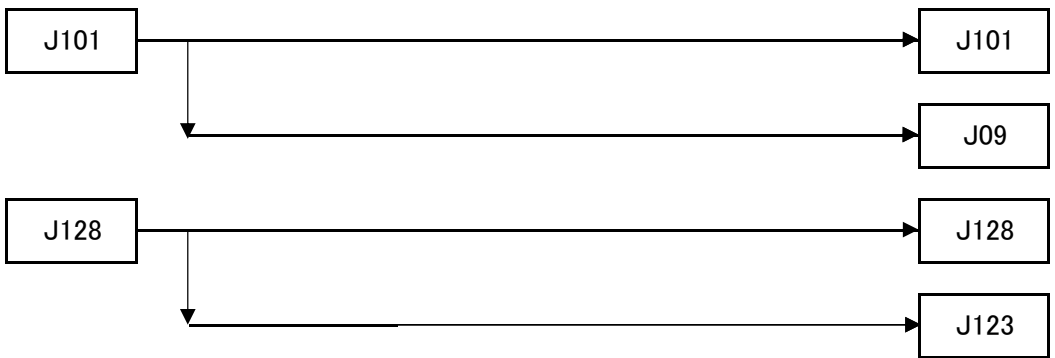
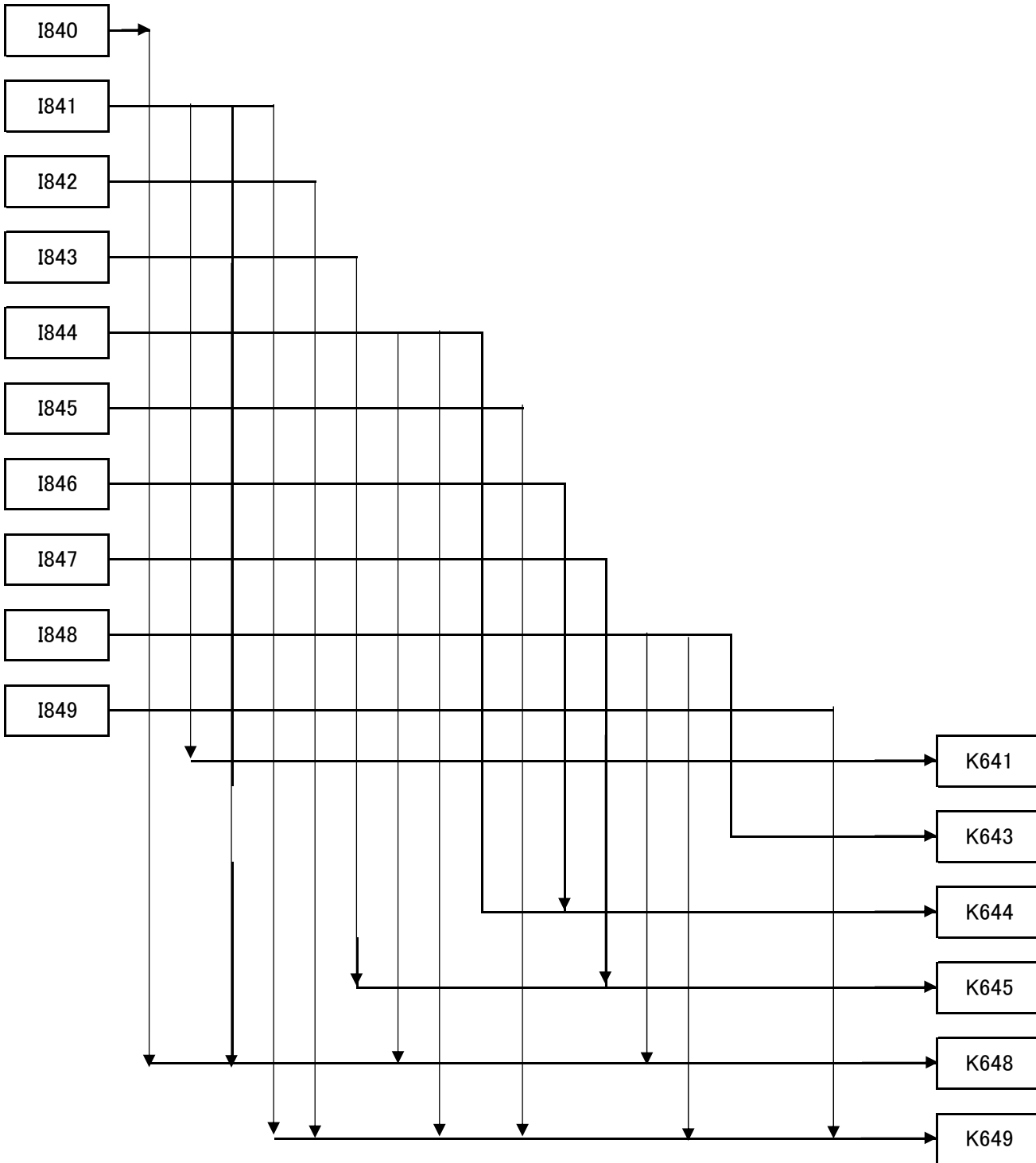


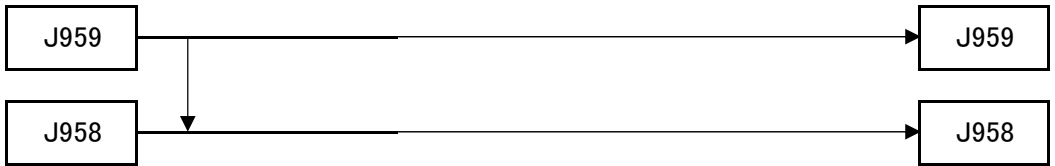
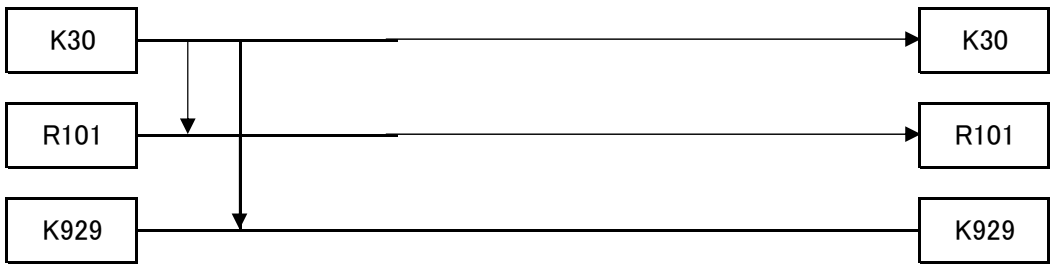
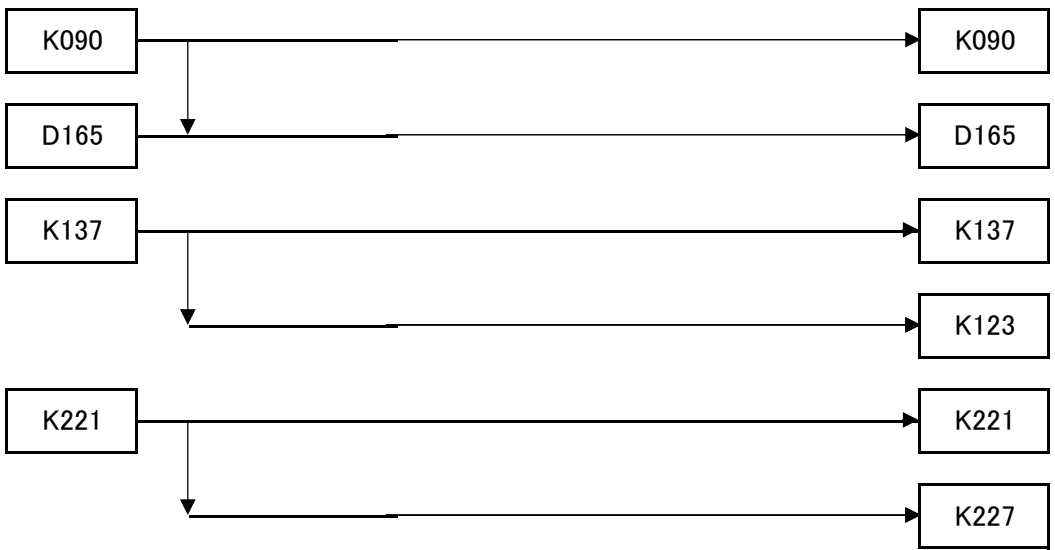
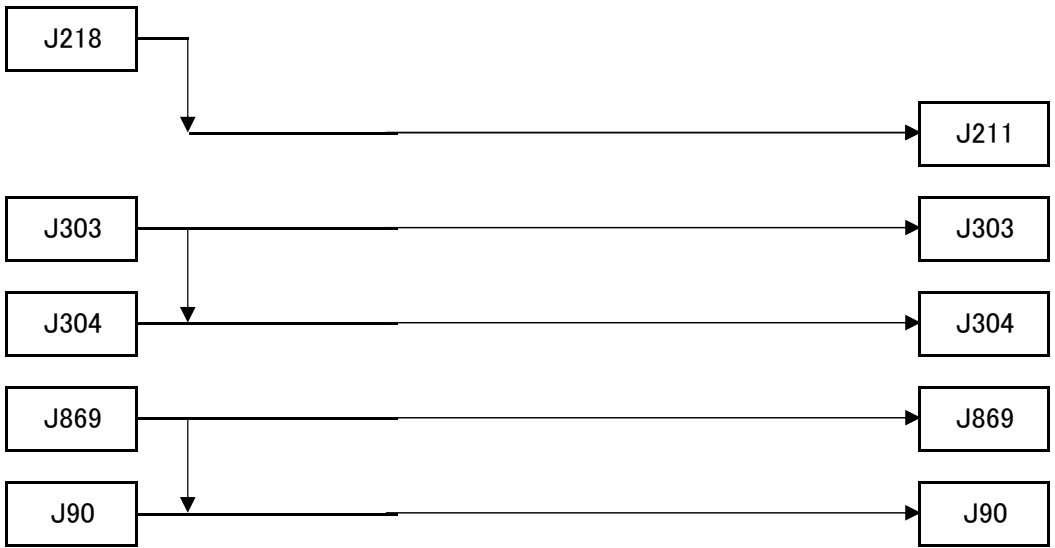


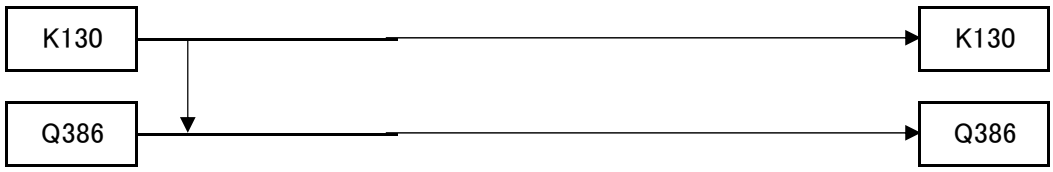


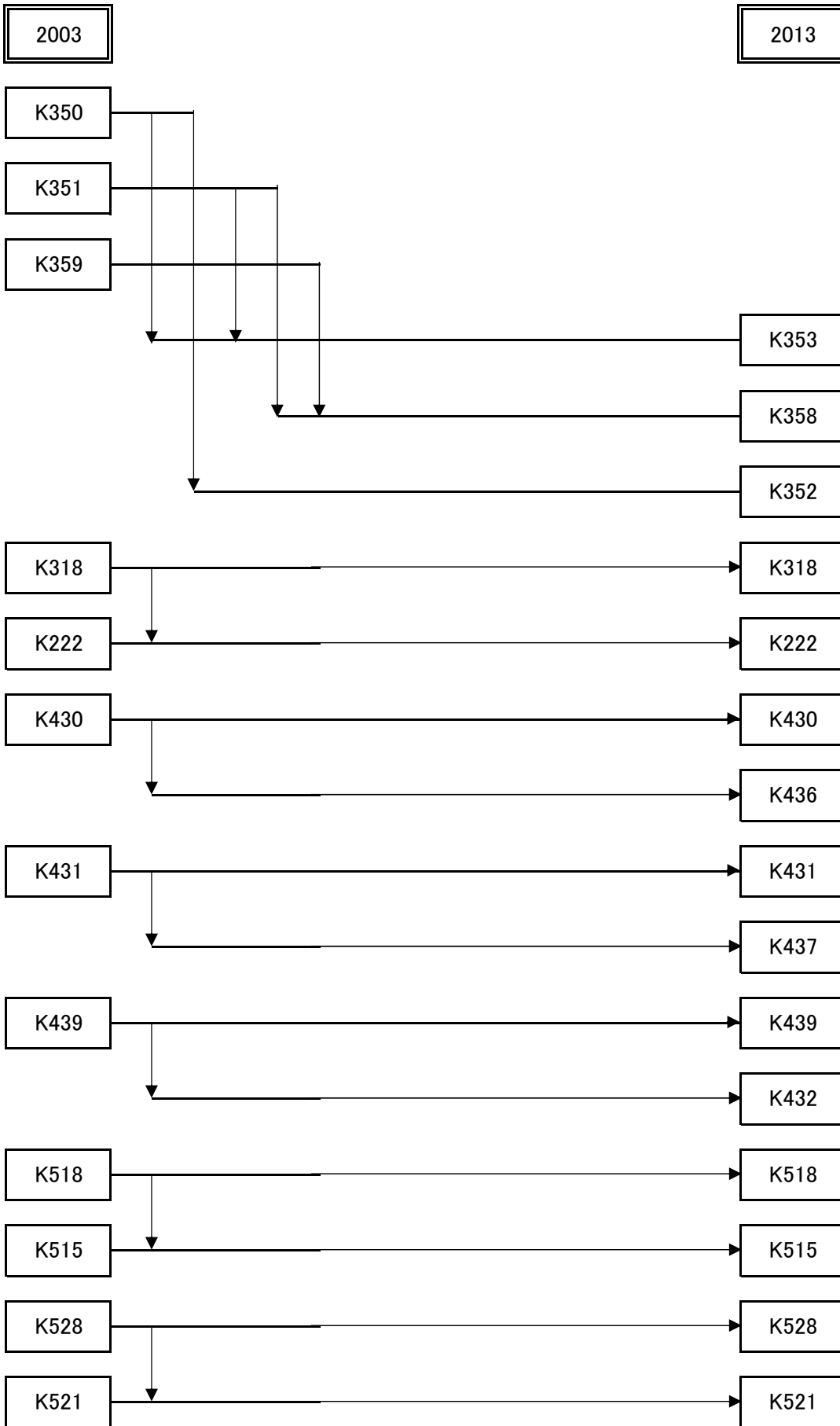
2003

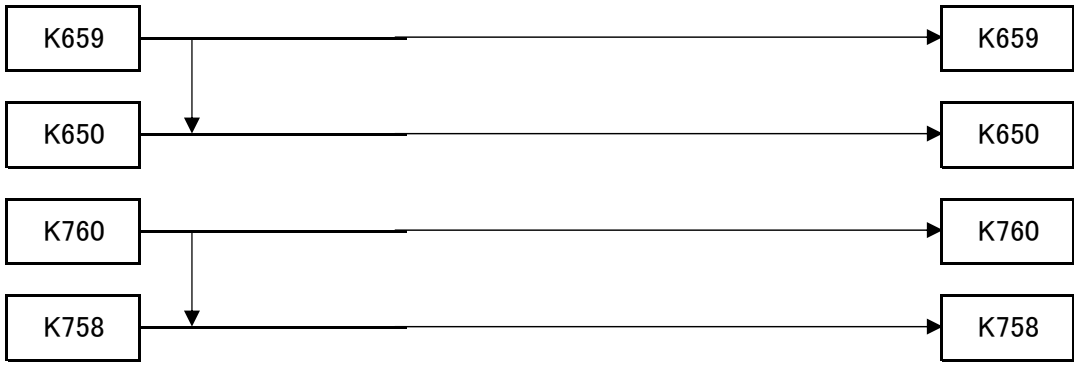
2013

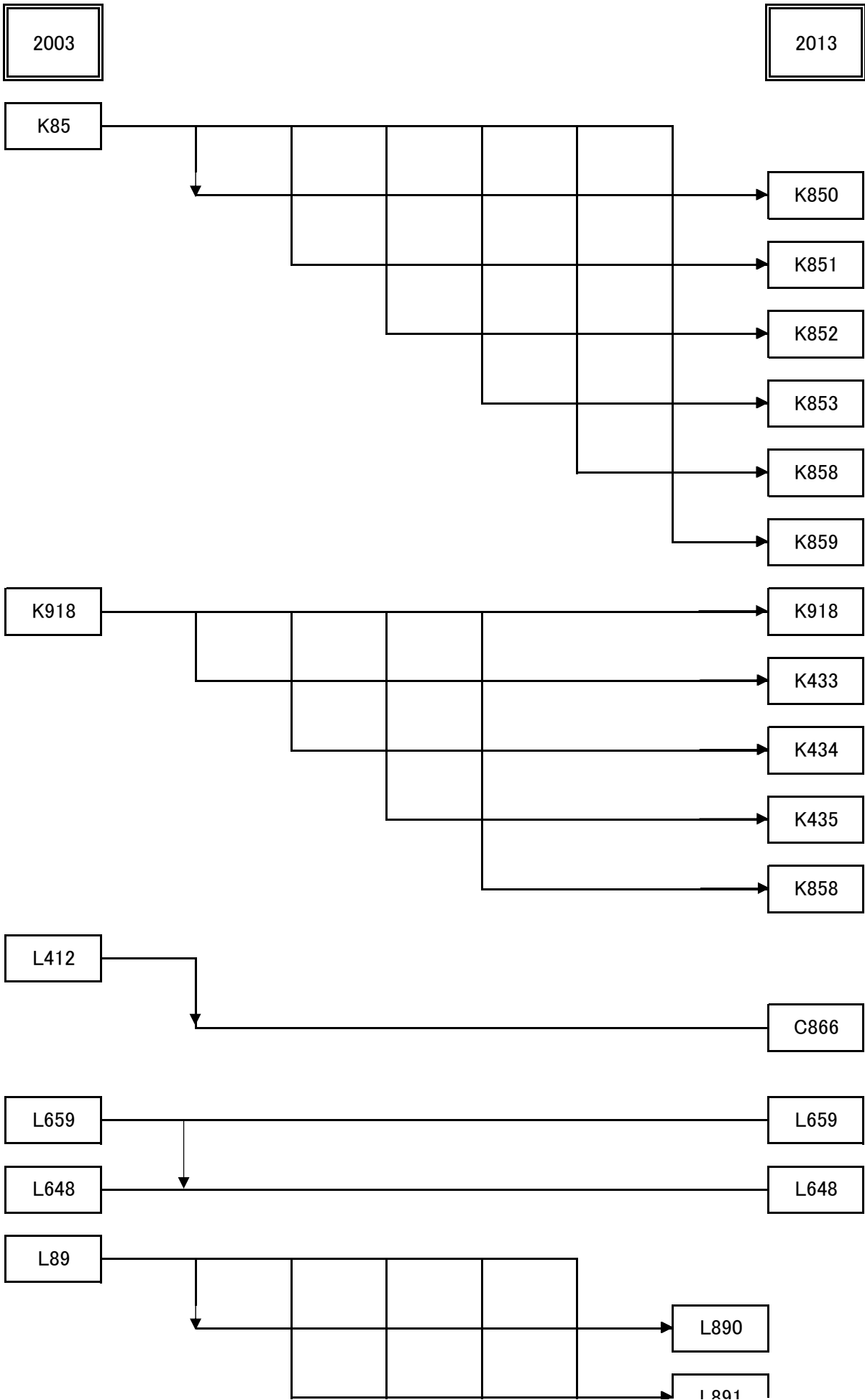


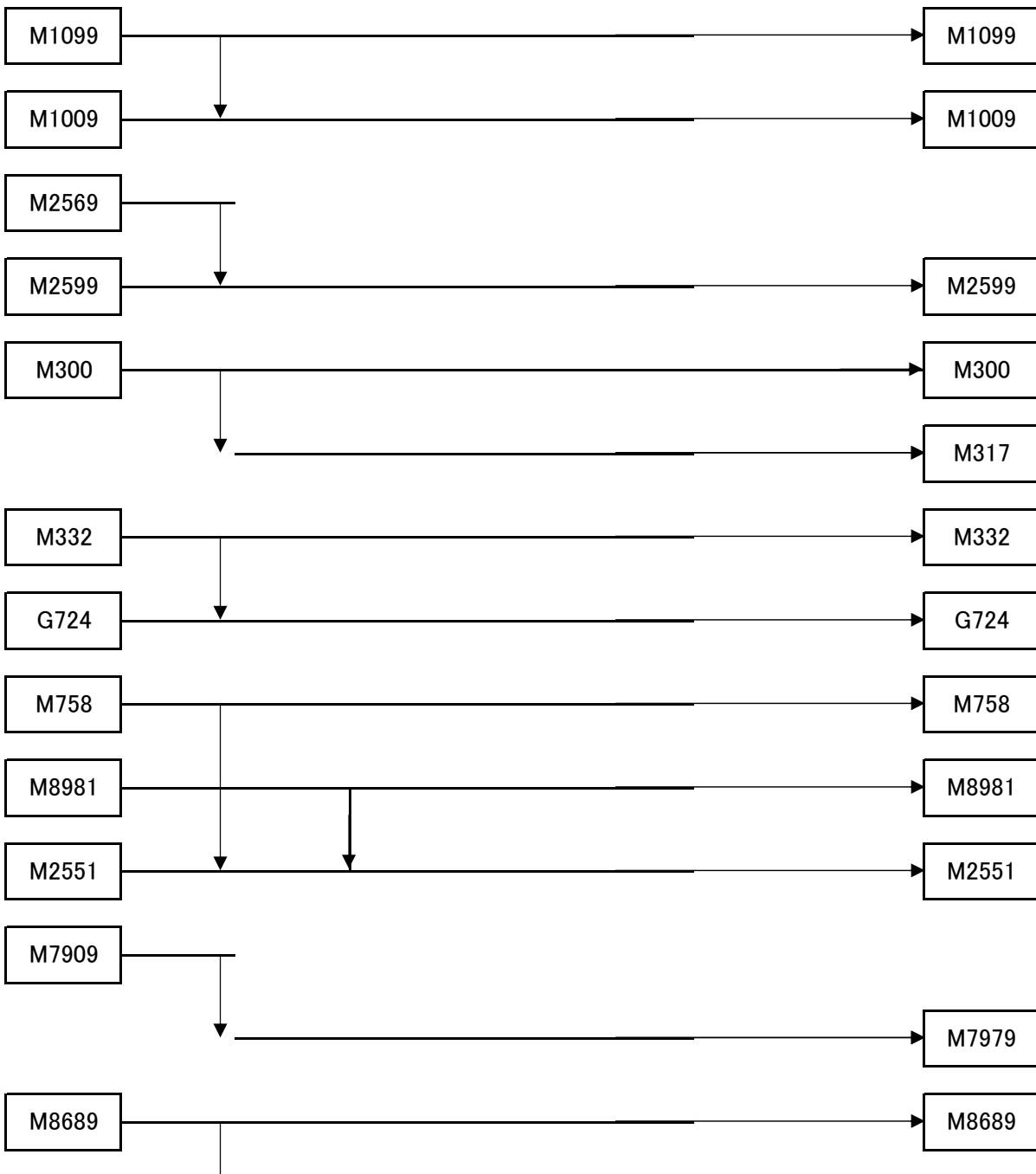
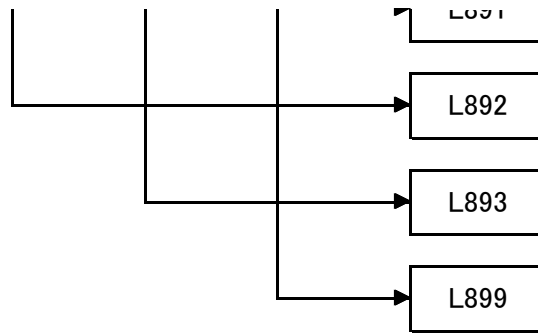


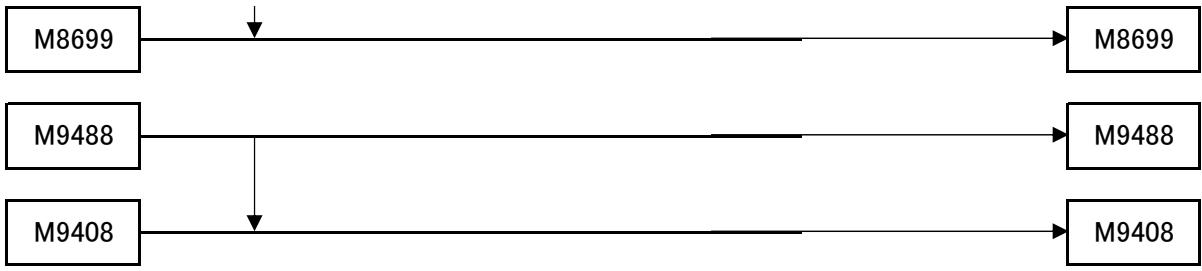


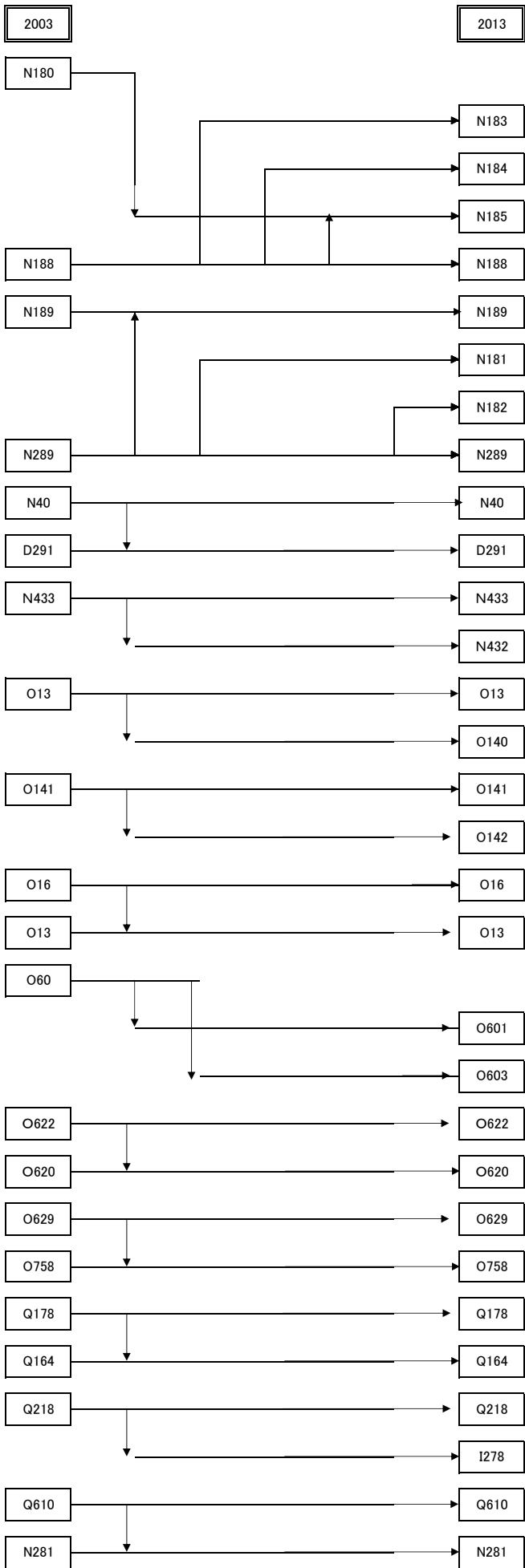


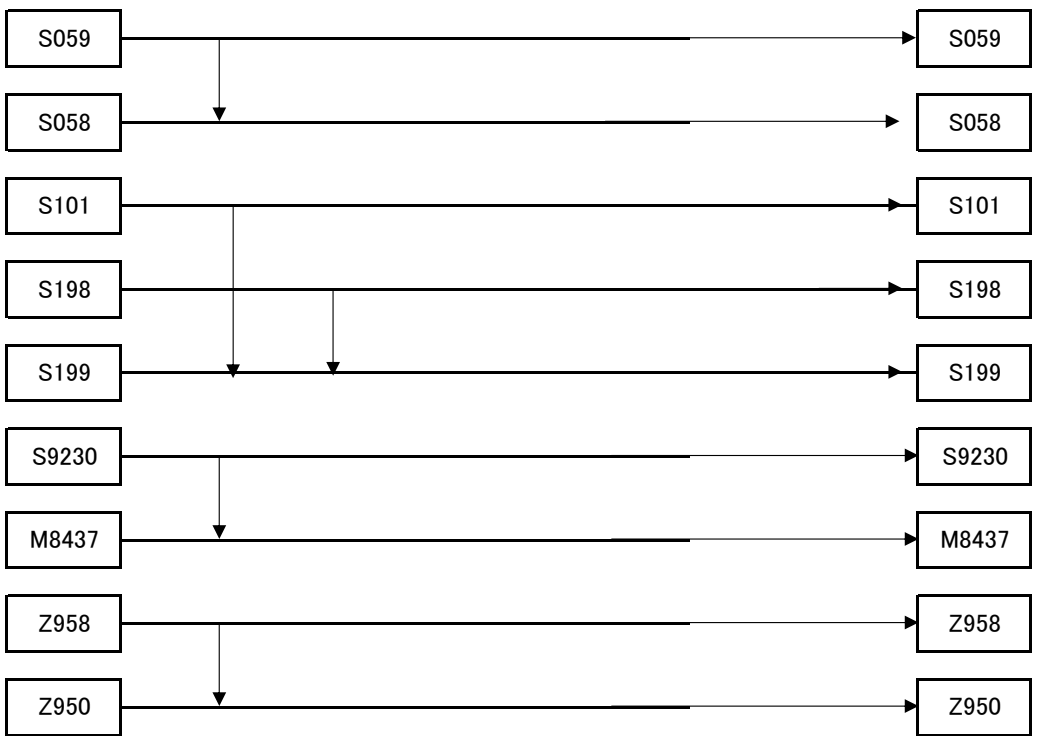
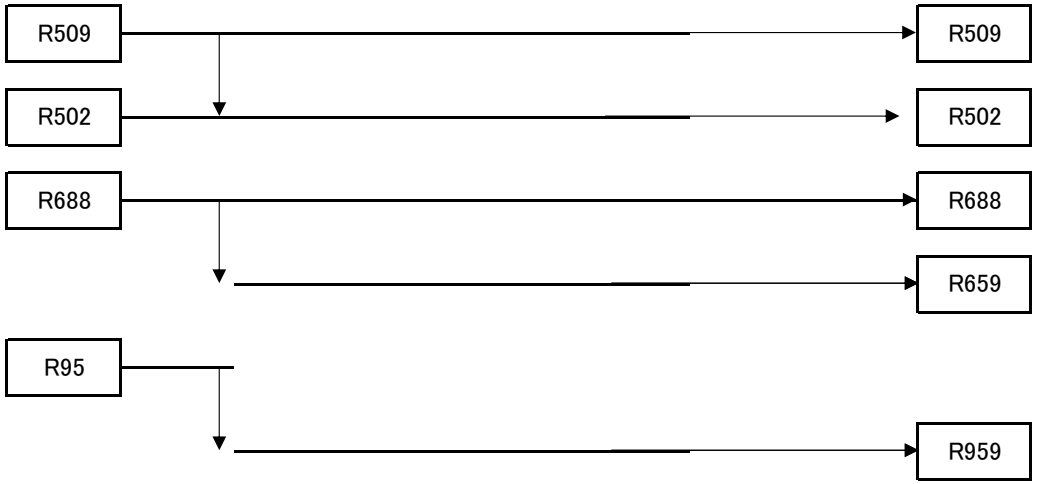
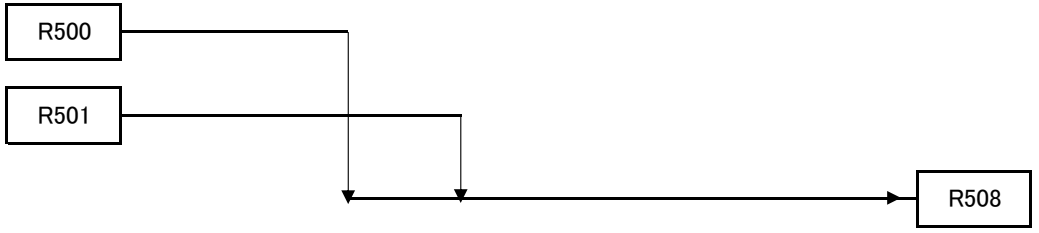




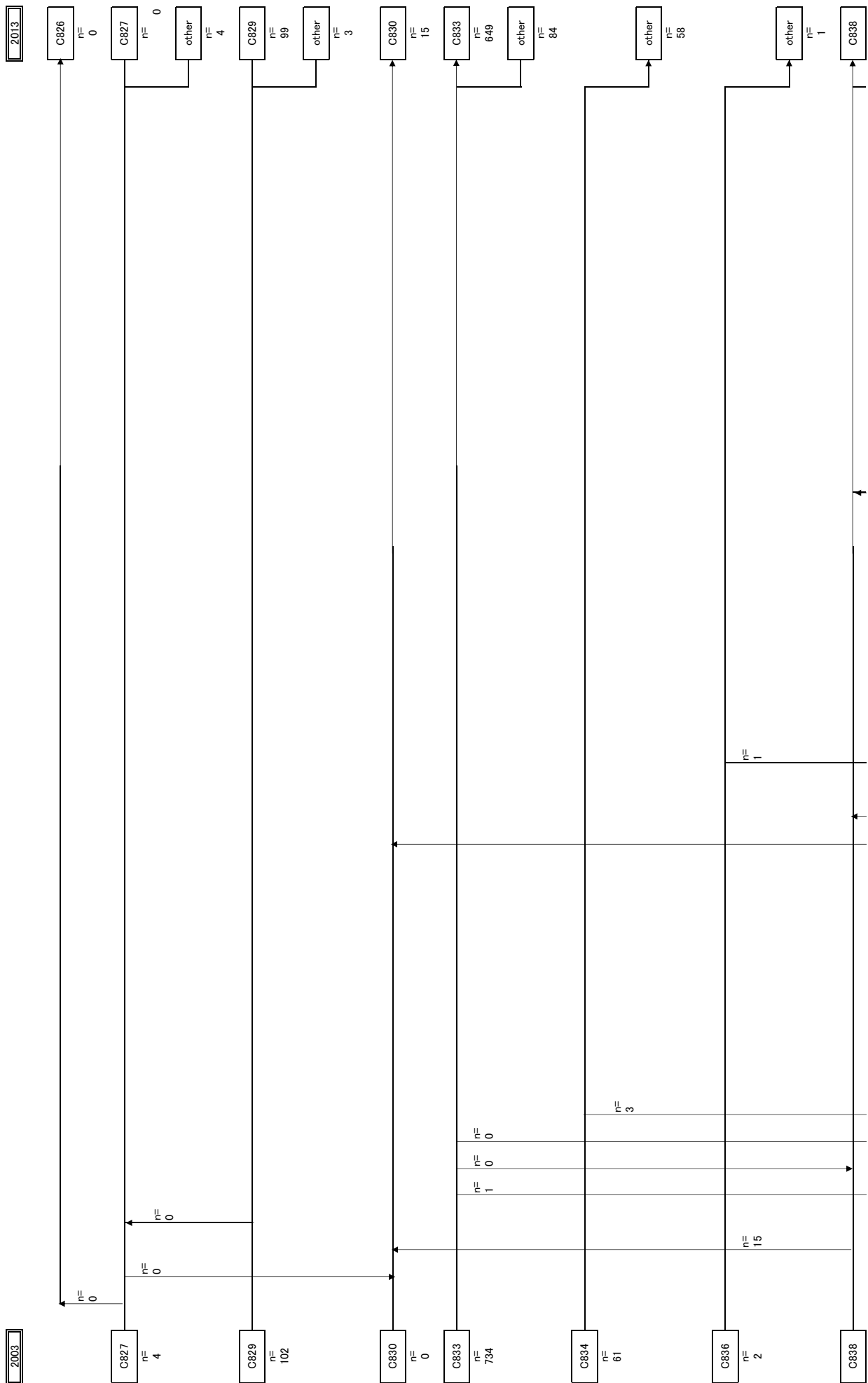


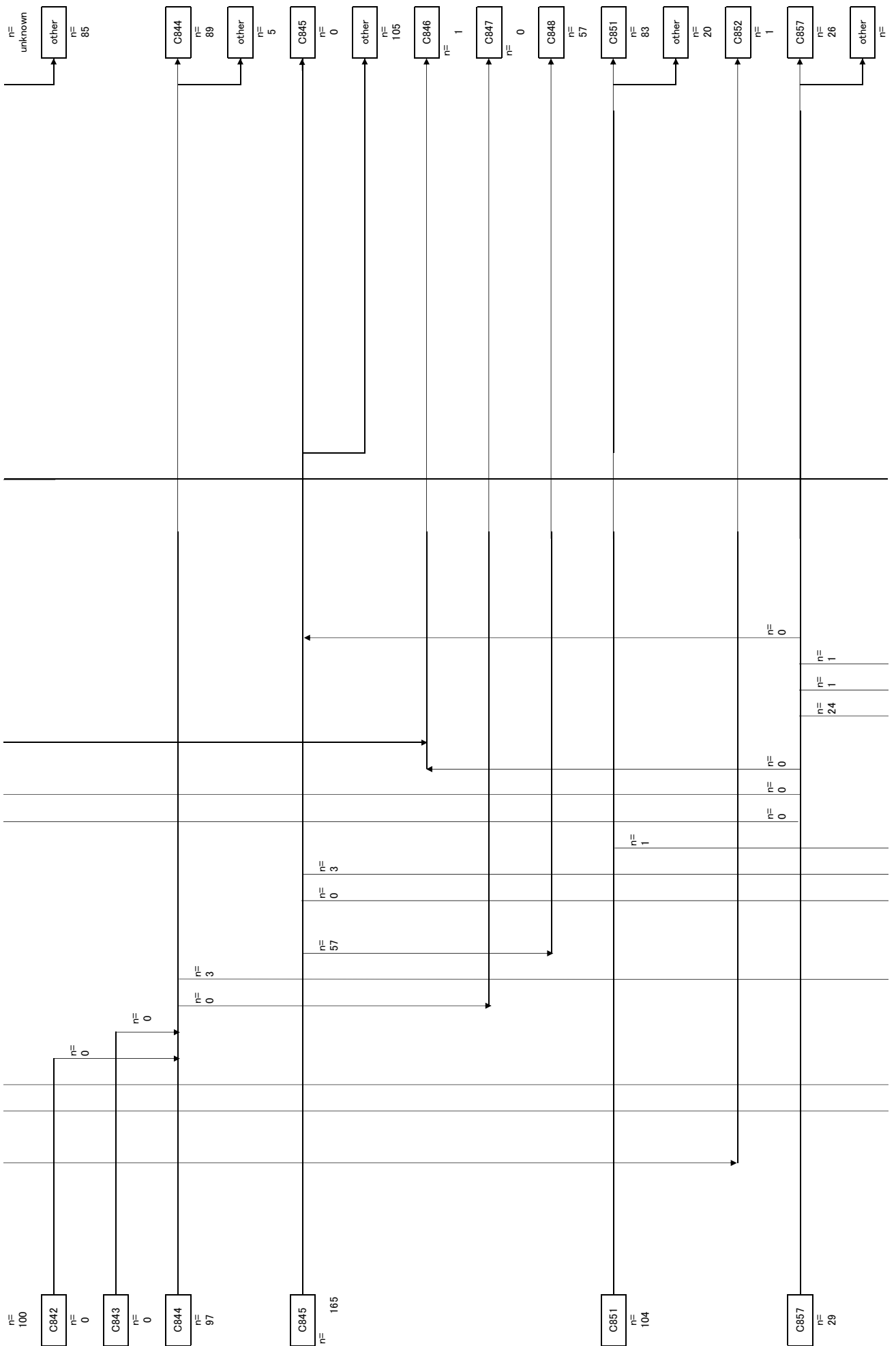


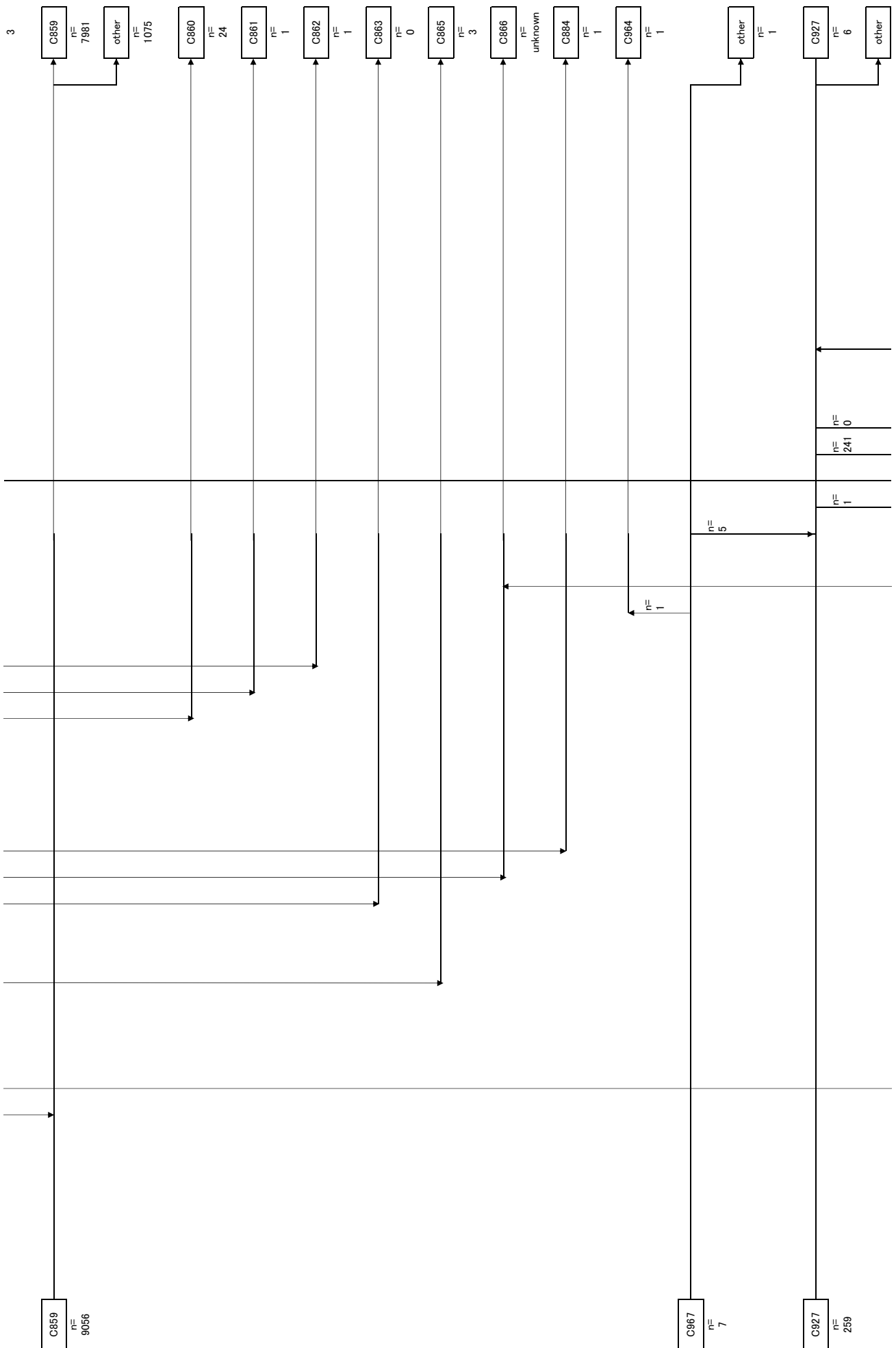


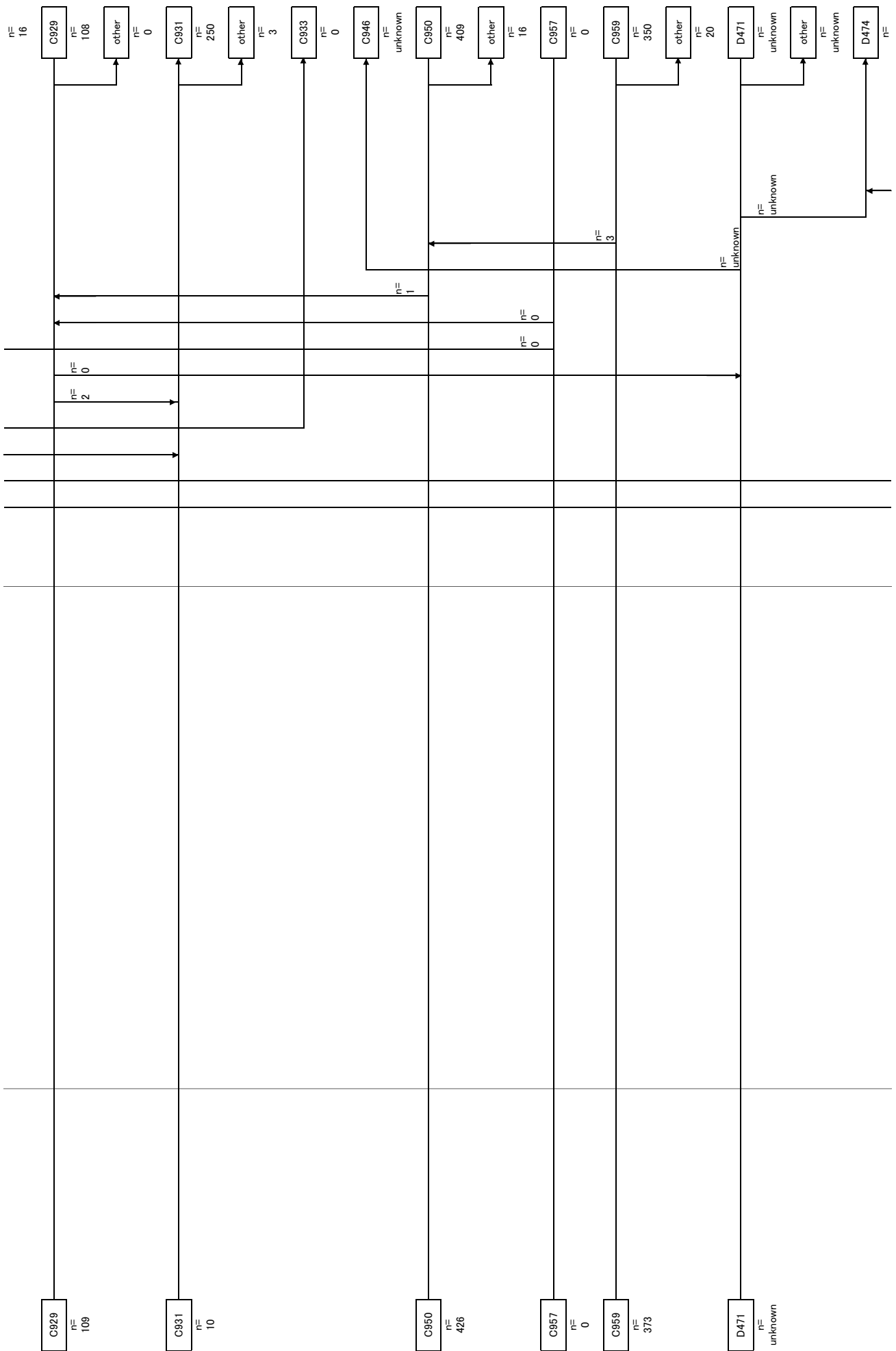


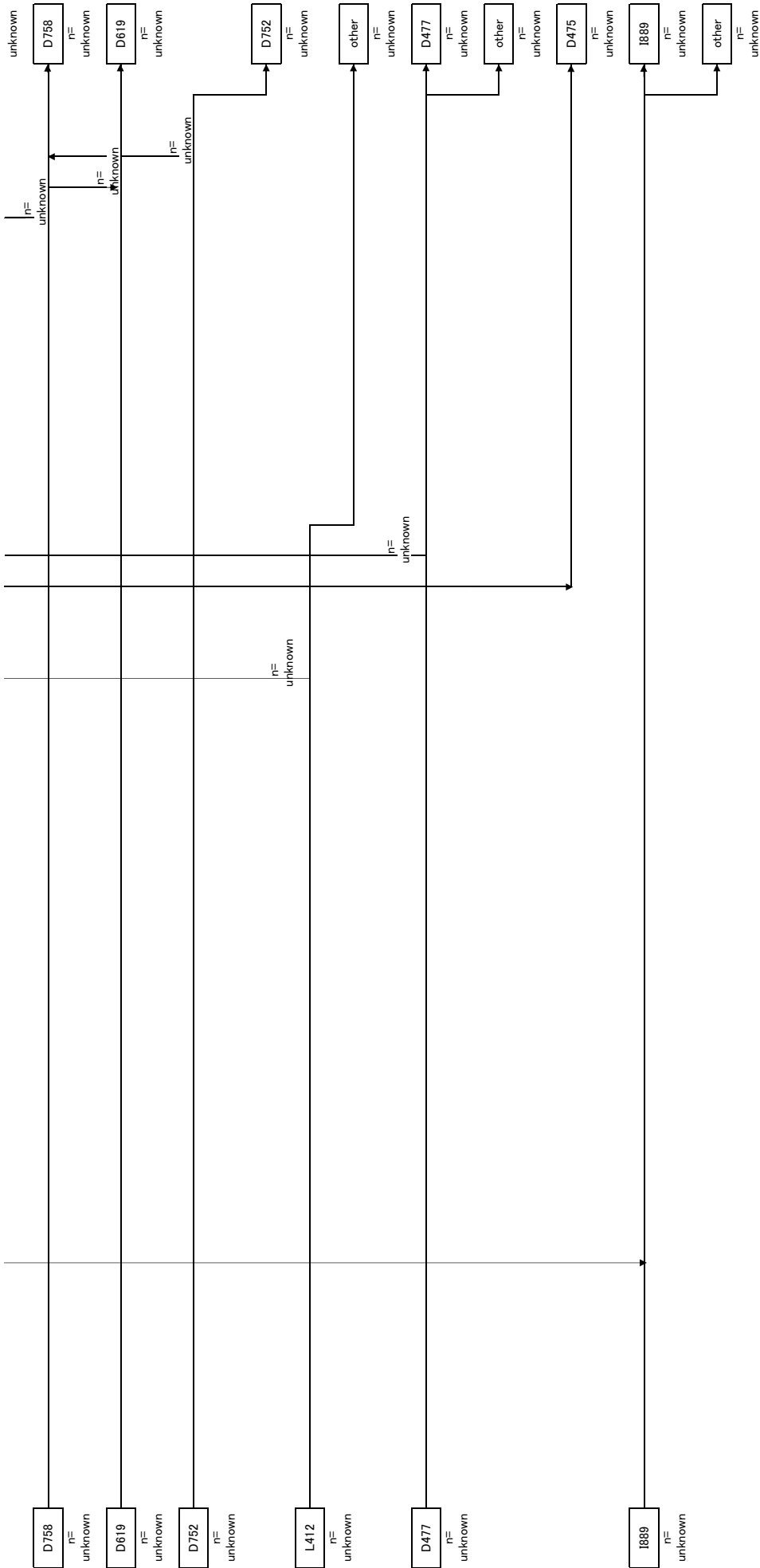
調整法(目視のみ)



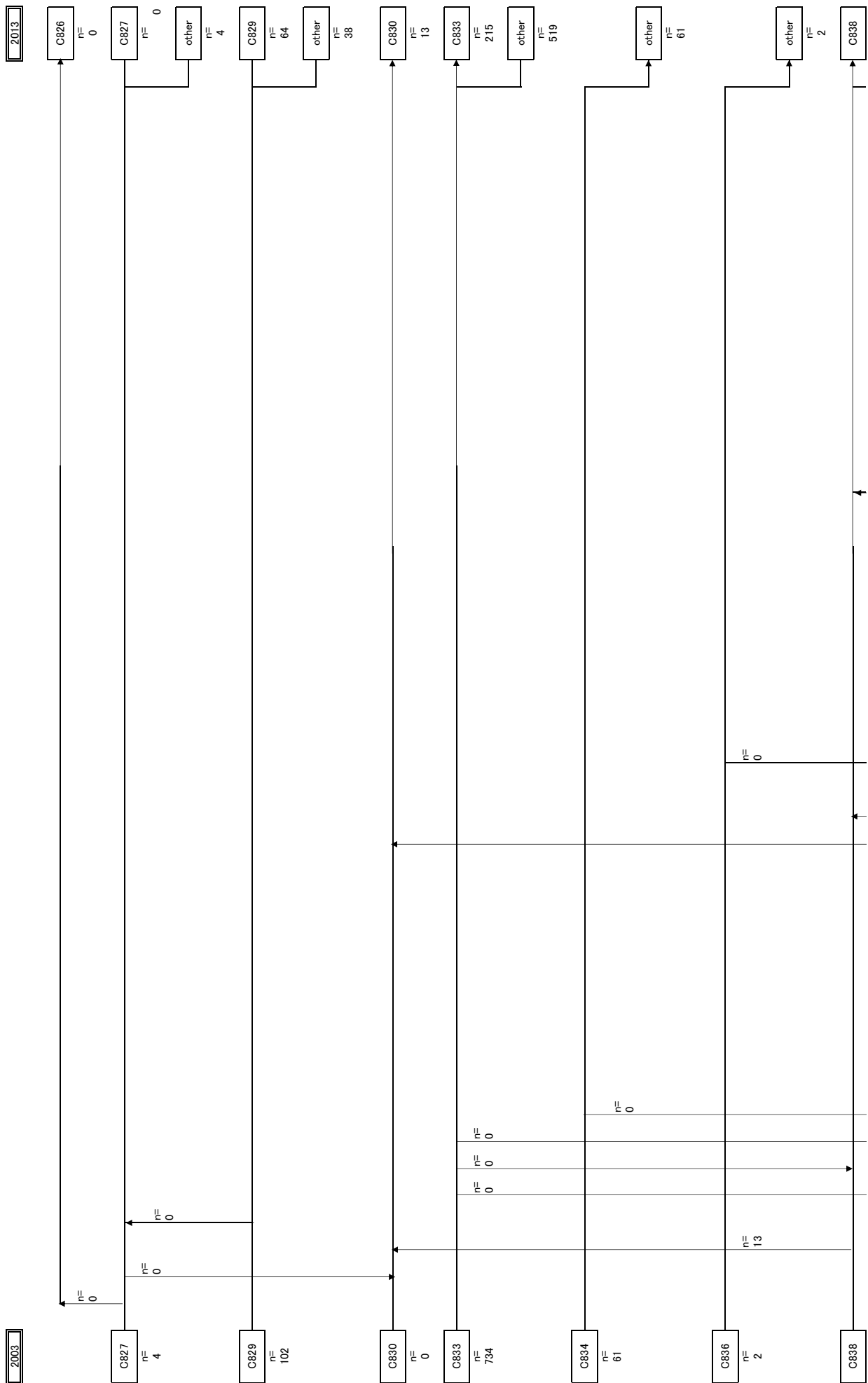


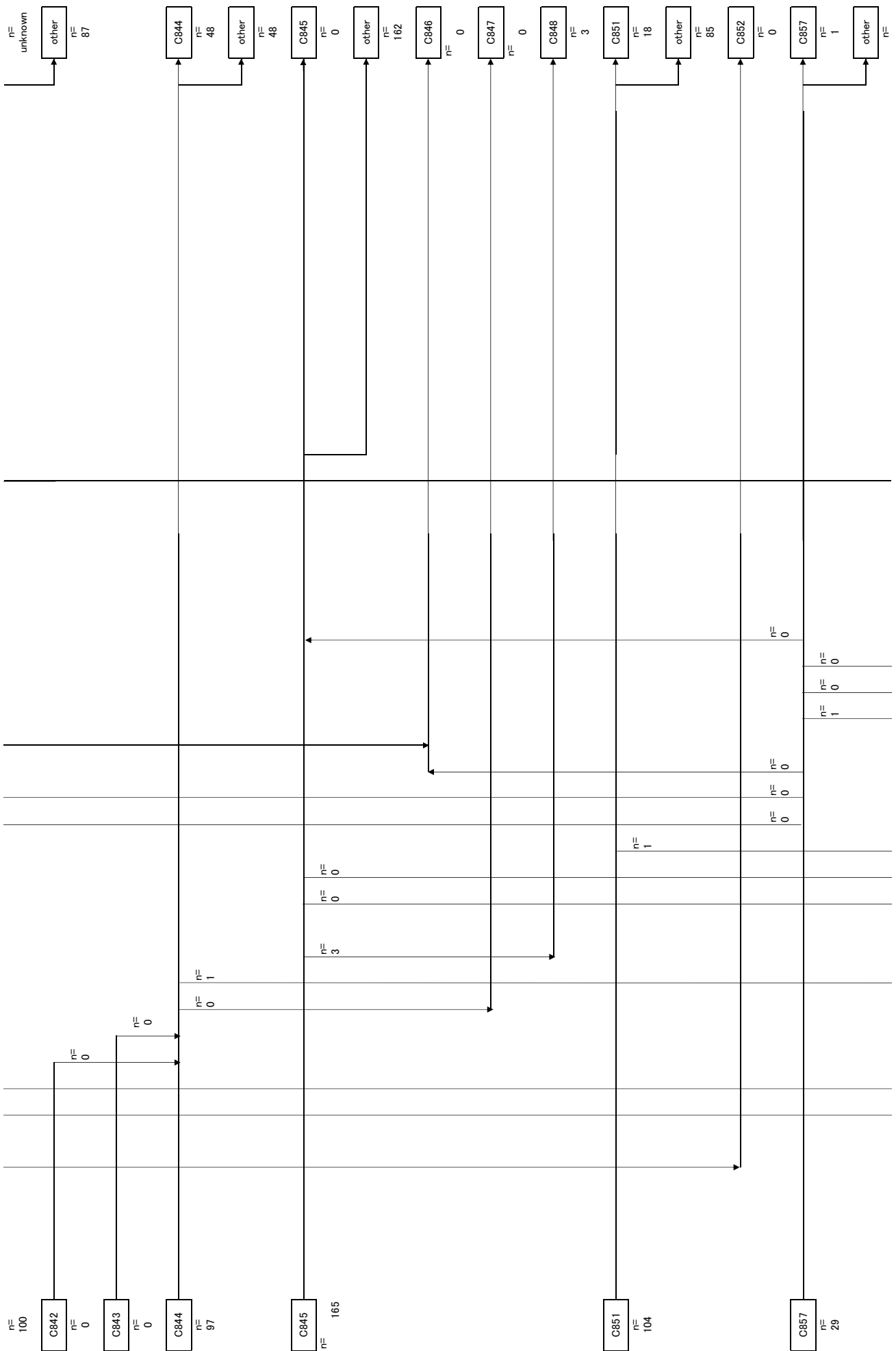


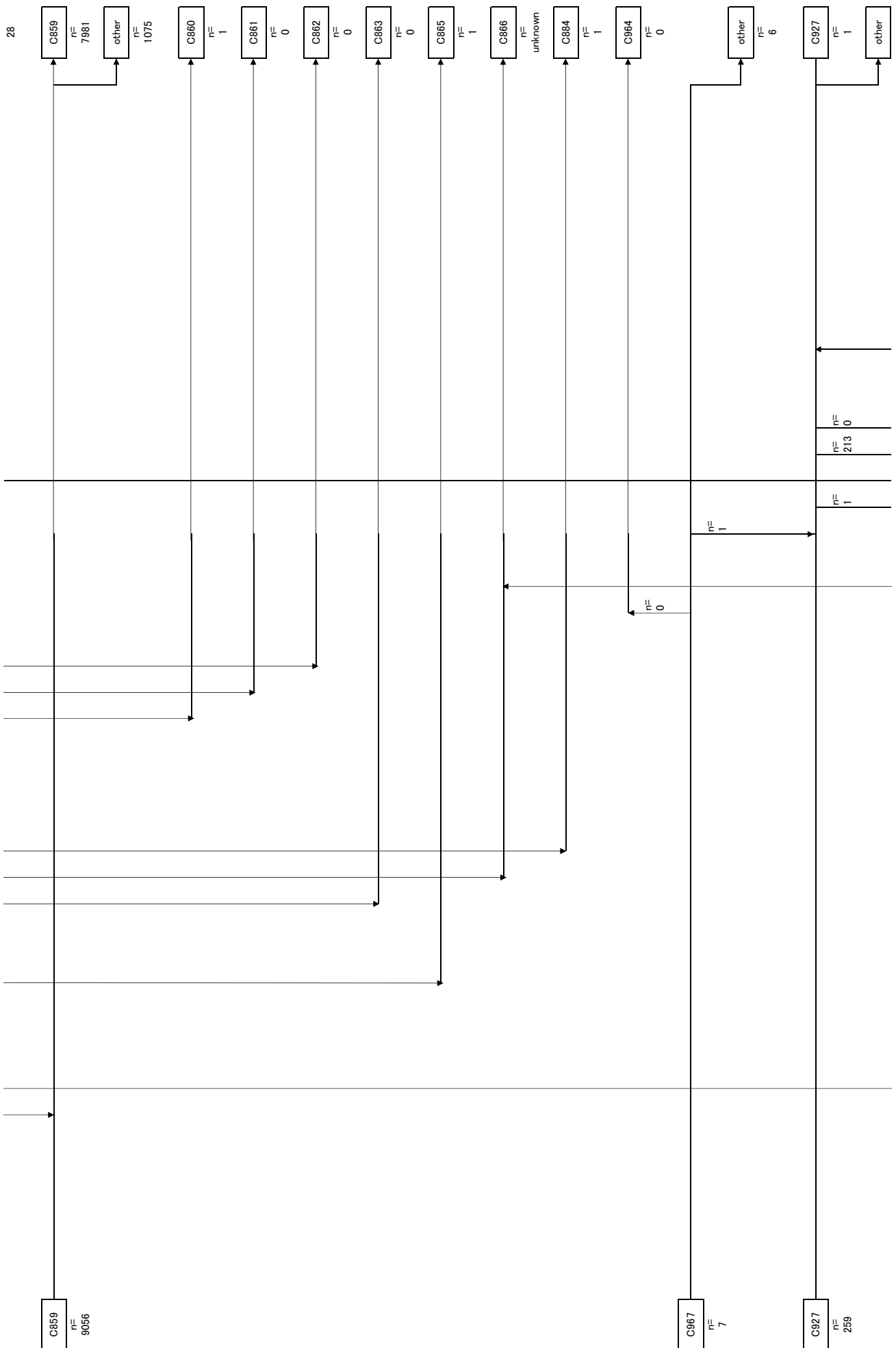


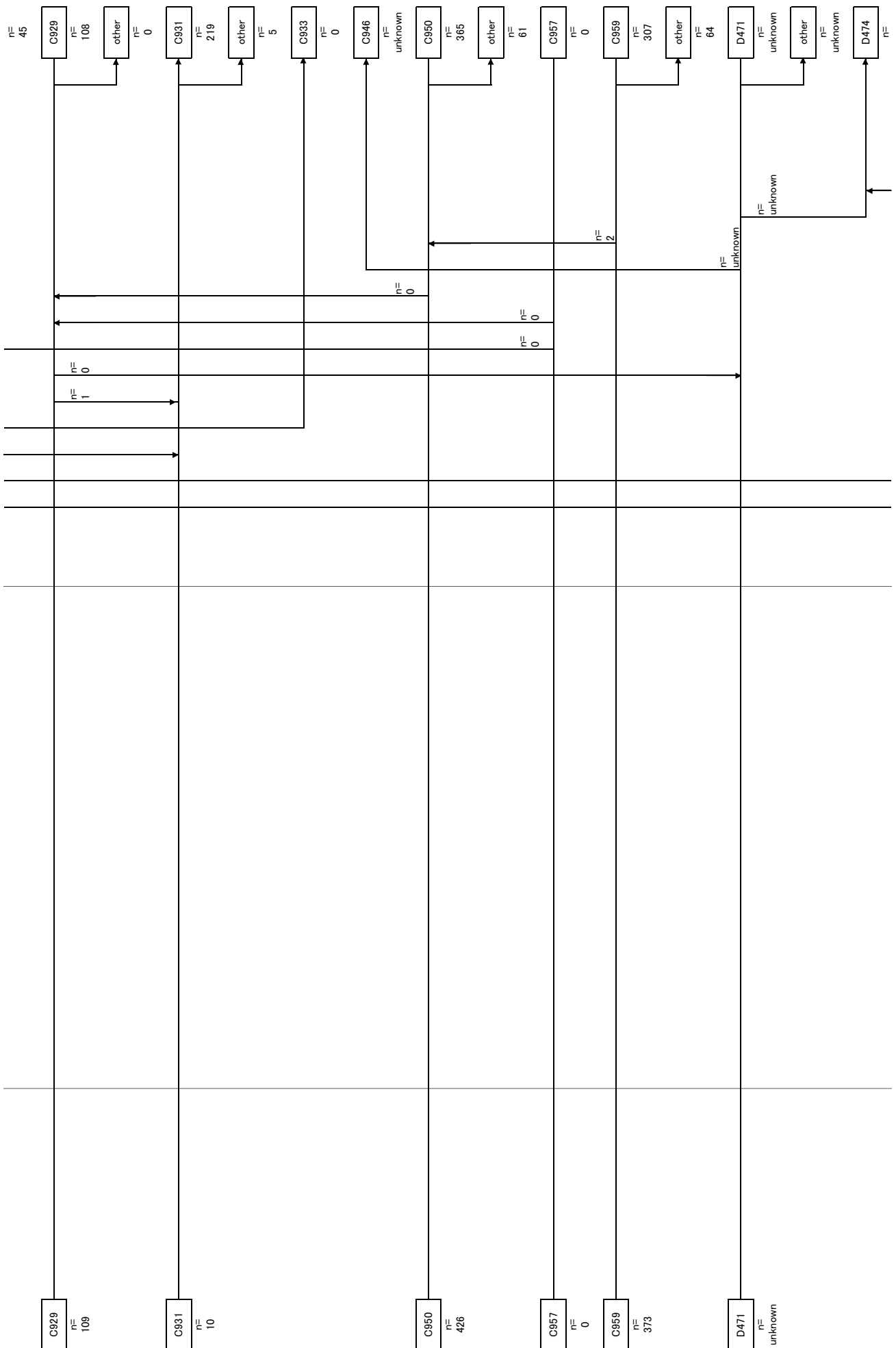


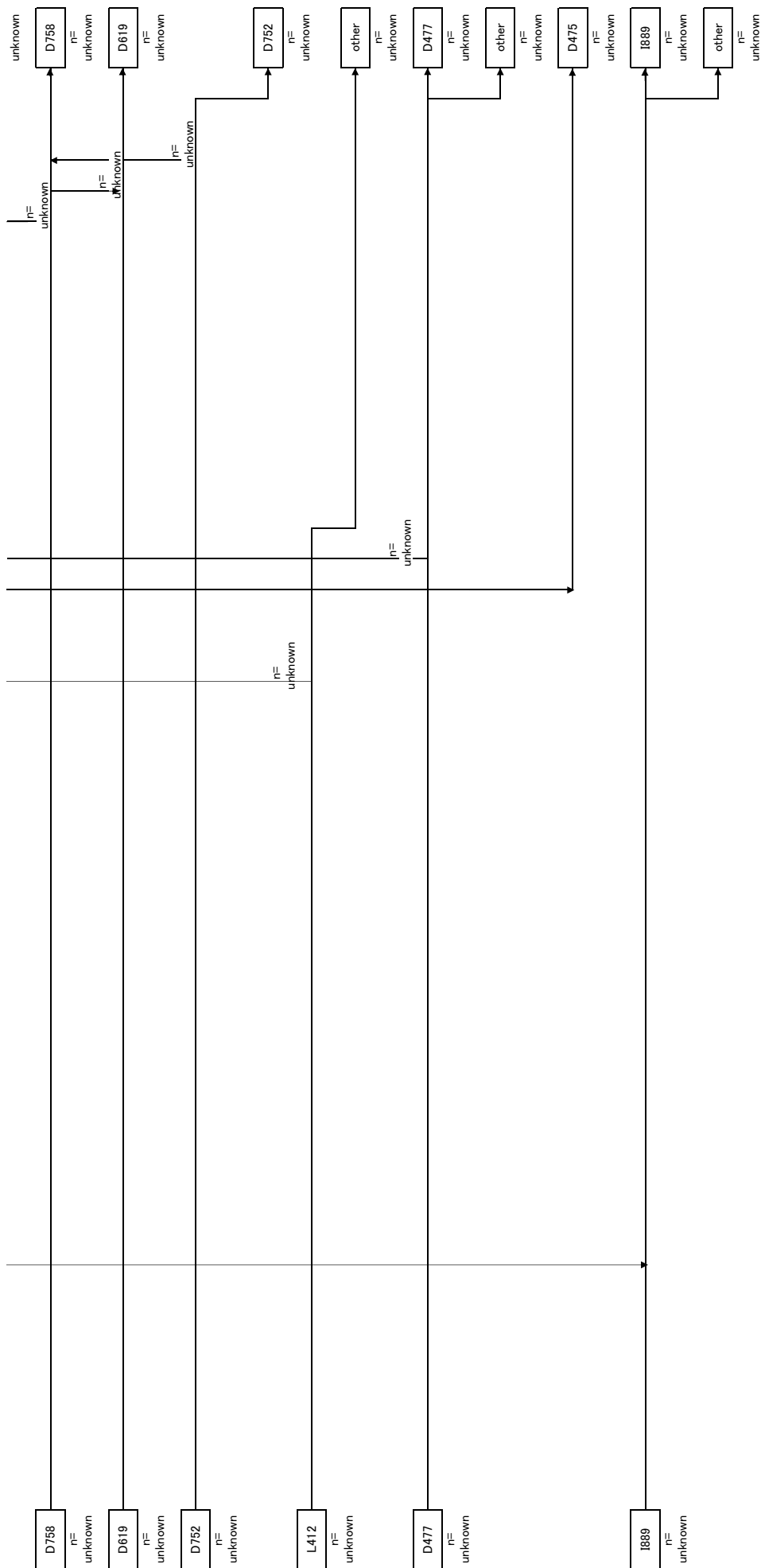
完全一致法







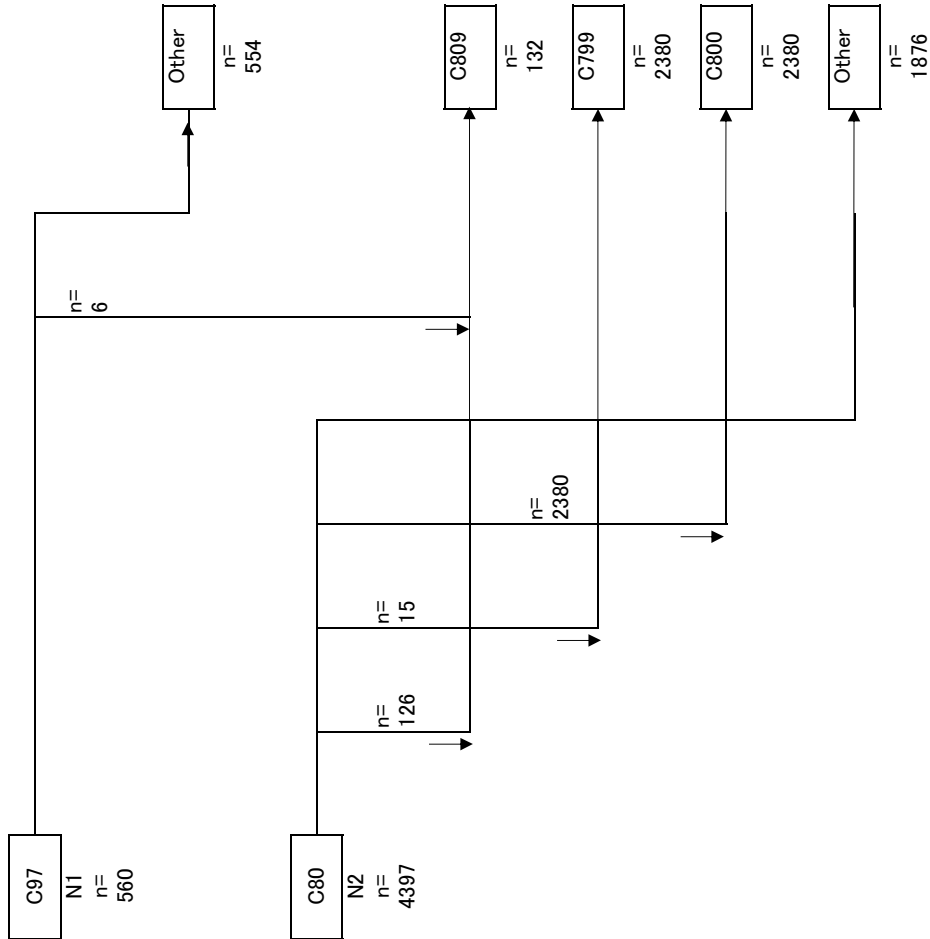




完全一致法

2003

2013

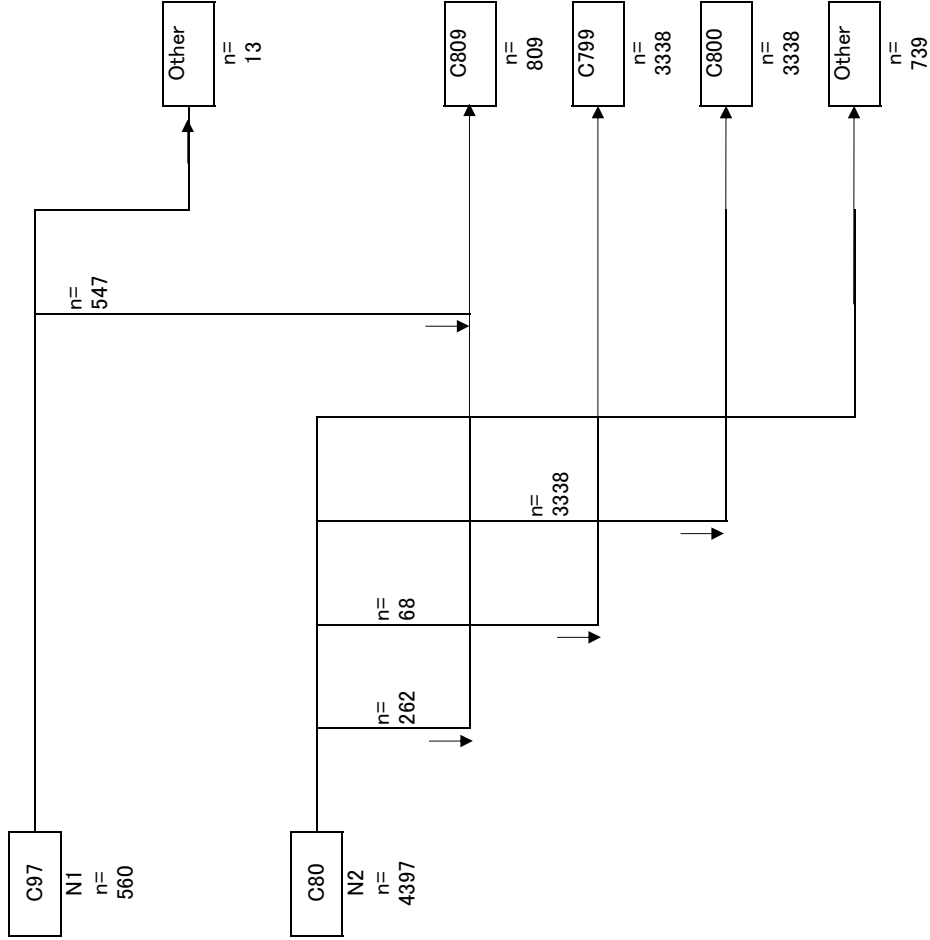


※本来C97、C80でコードされるべき症例は加味していない

調整法(目視のみ・再コーディング困難)

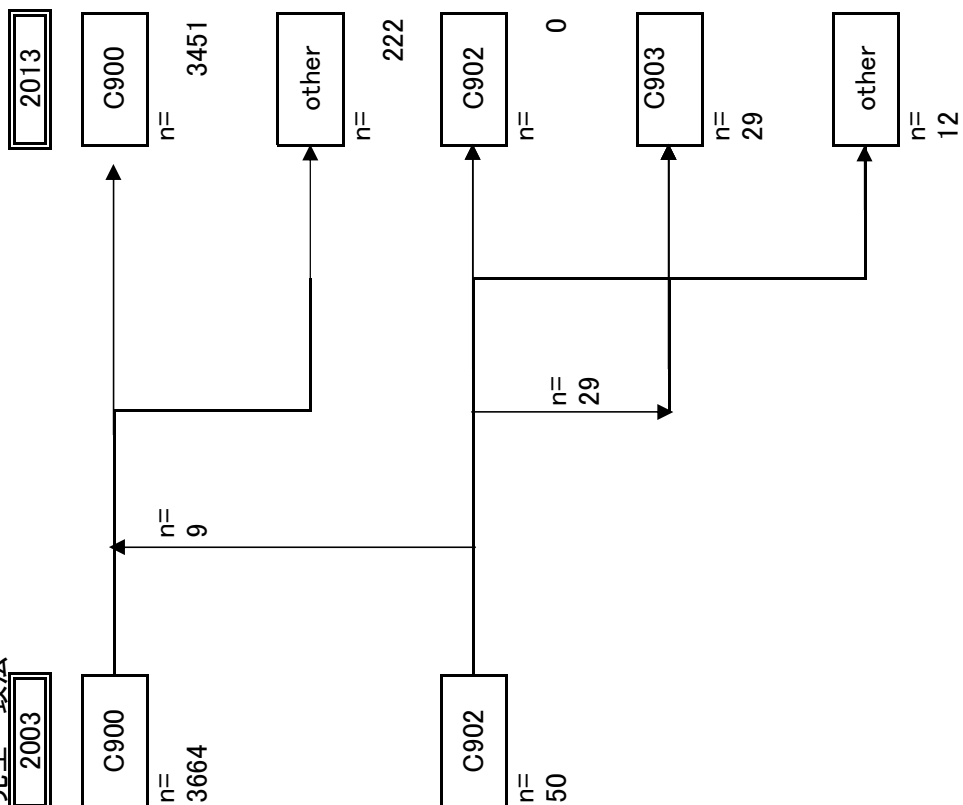
2003

2013

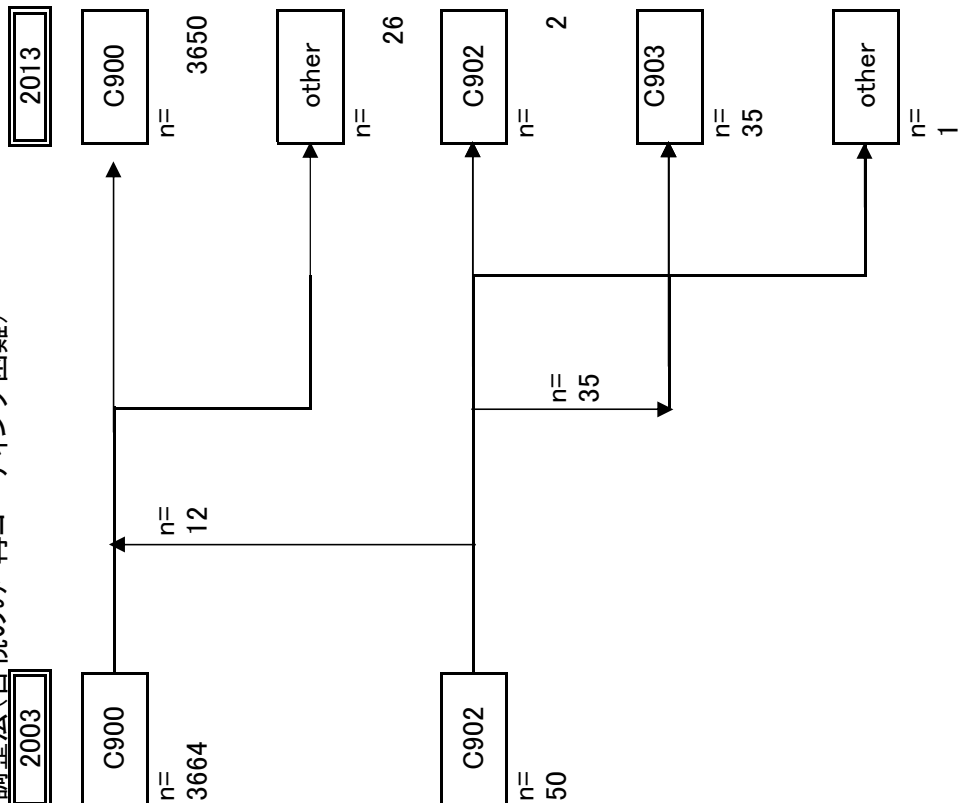


※本来C97、C80でコードされるべき症例は加味していない

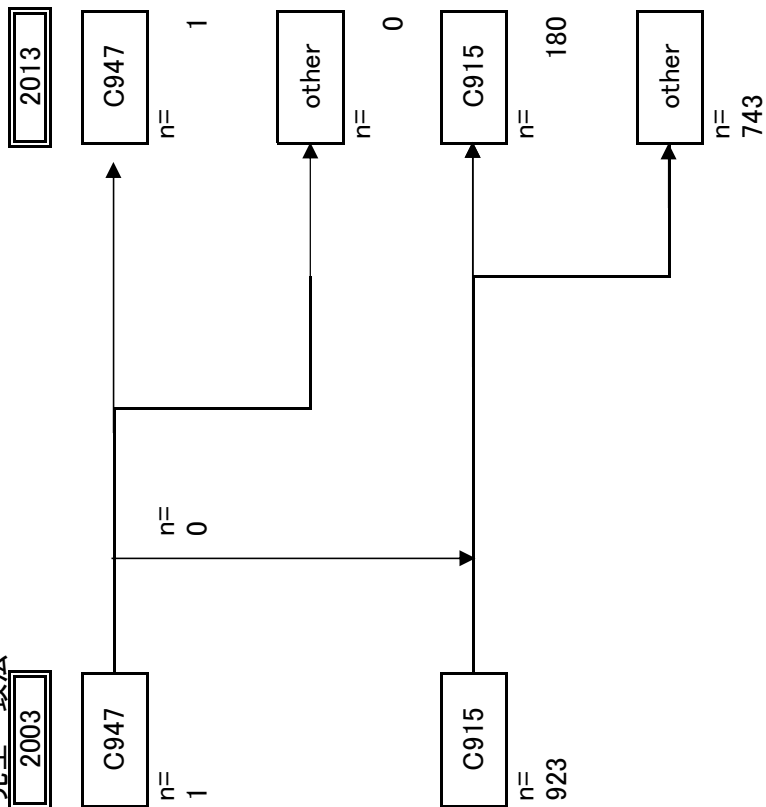
完全一致法



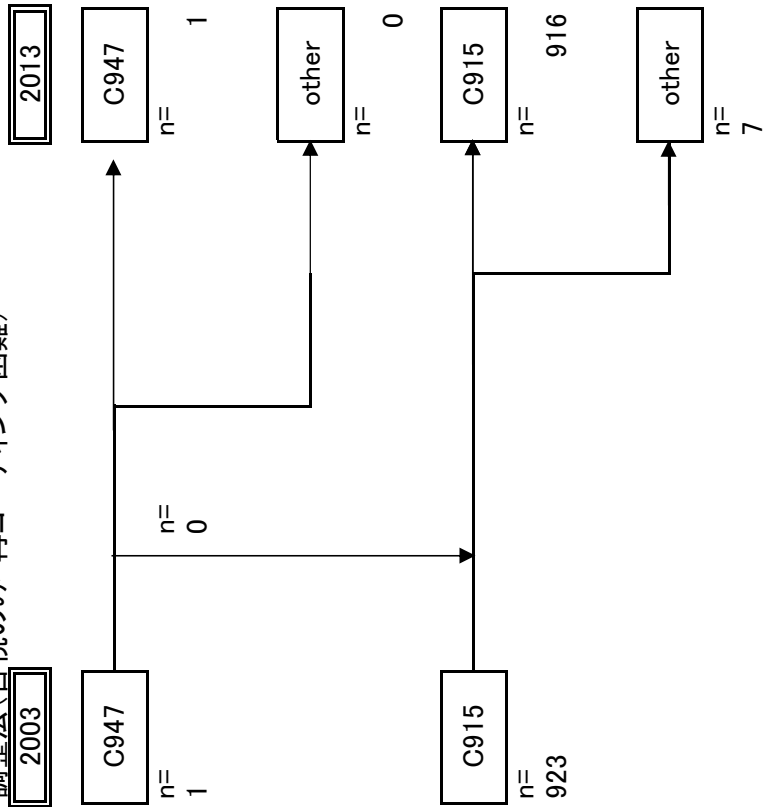
調整法(目視のみ・再コーディング困難)



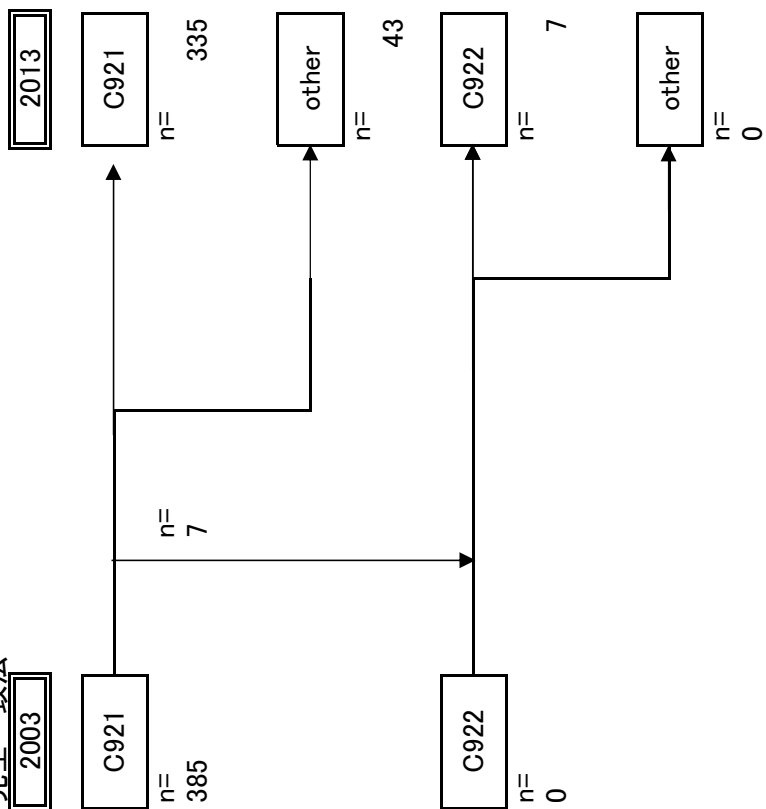
完全一致法



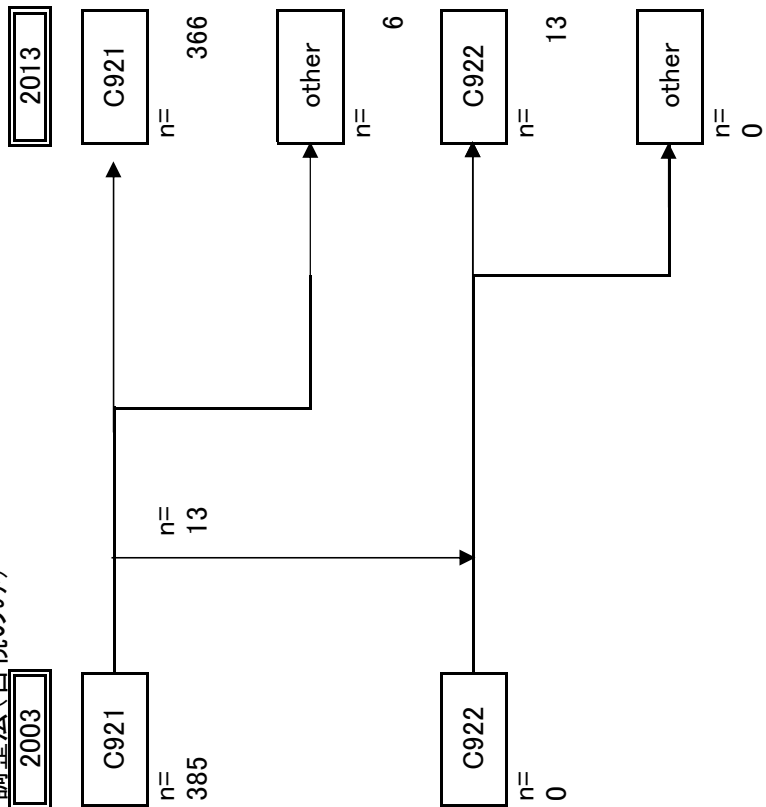
調整法(目視のみ・再コーディング困難)



完全一致法



調整法(目視のみ)



完全一致法

2003

C961
n= 29

n= 5

other
n= 24

C968
n= 5

2013

調整法(目視のみ・再コーディング困難)

2003

C961
n= 29

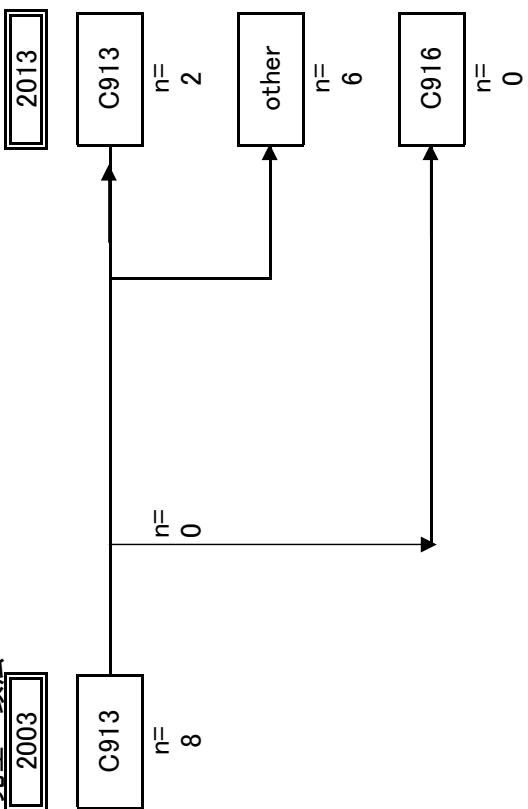
n= 28

other
n= 1

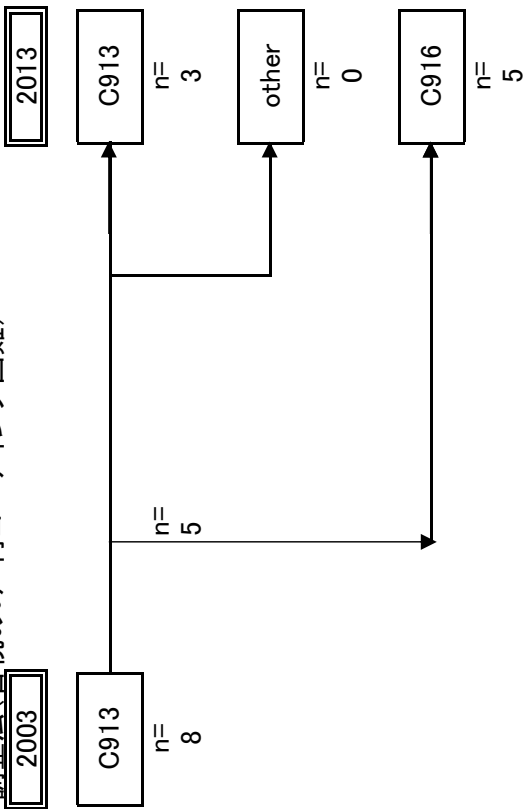
C968
n= 28

2013

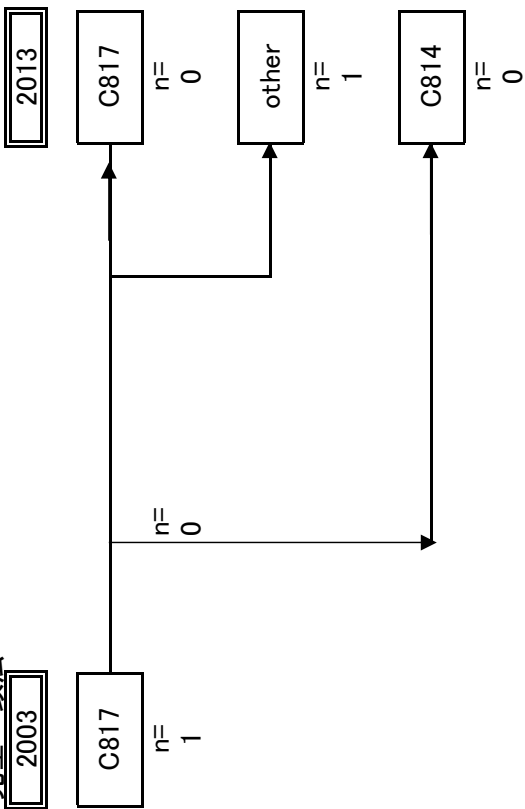
完全一致法



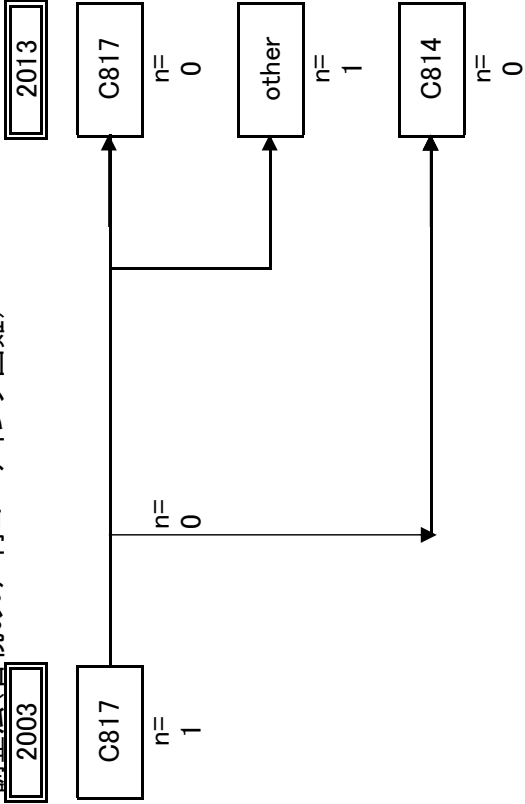
調整法(目視のみ・再コーディング困難)



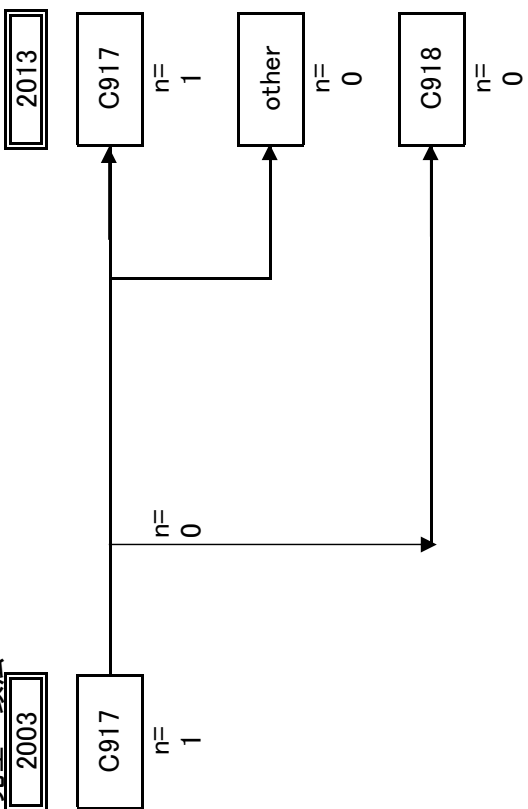
完全一致法



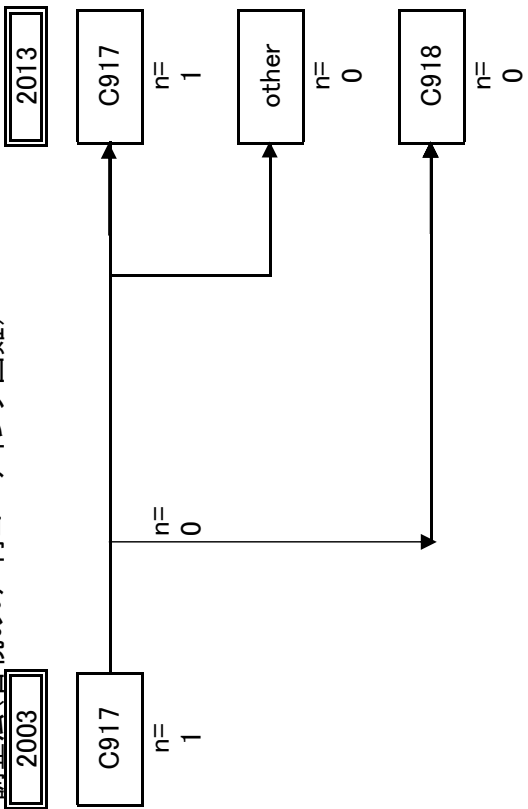
調整法(目視のみ・再コーディング困難)



完全一致法



調整法(目視のみ・再コーディング困難)



研究成果の刊行に関する一覧表

なし