

フォント情報交換国際規格に 関する最新動向とその課題

広島大学総合科学研究科 鈴木俊哉
大阪工業大学情報科学部 小町祐史

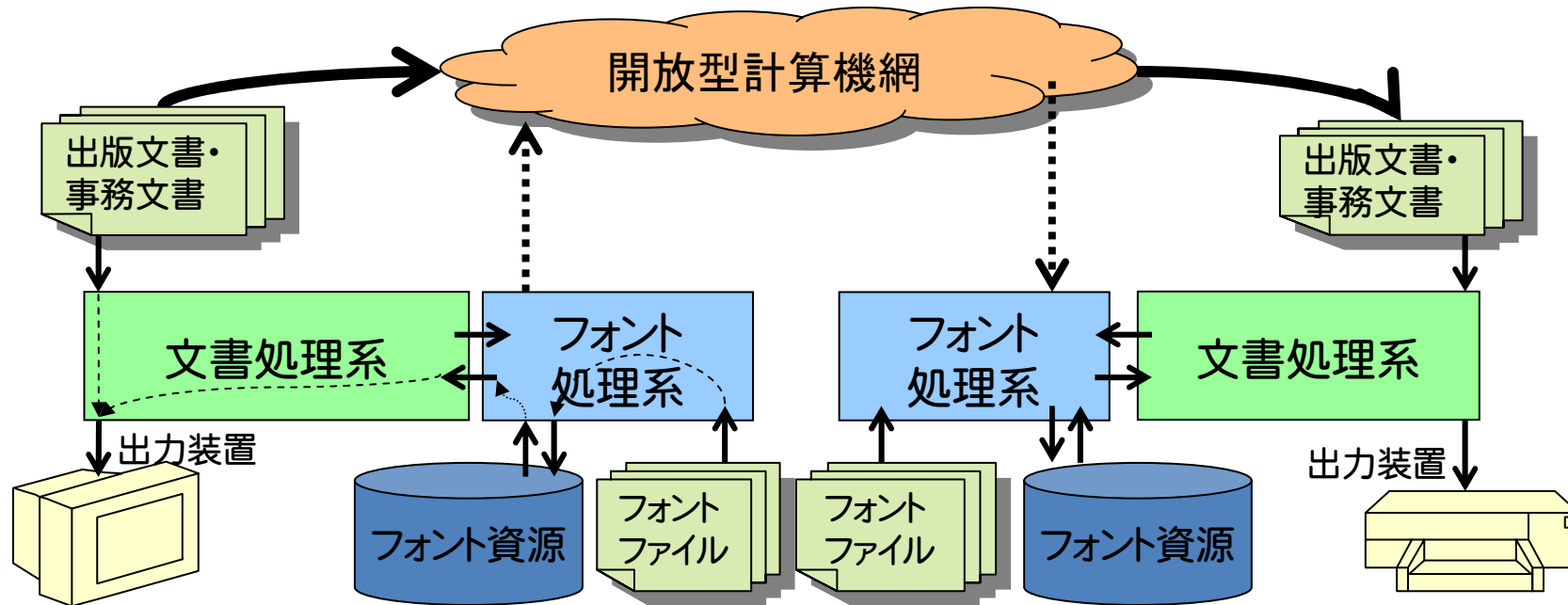
発表の構成

- 背景
 - ISO/IEC 9541
 - ✓ ISO/IEC 9541規格群
 - ✓ ISO/IEC 9541のスコープ
 - フォントのデータ構造を規定する規格
- UnicodeテキストとOpenType
 - 従来の欧文用のフォントの構造
 - 複雑な書字モデルの例(クメール文字)
 - ✓ Unicodeテキスト描画処理の流れ
 - ISO/IEC 14496-22規格
 - ✓ SC29/WG11によるOpenType標準化
 - ✓ OpenTypeを用いたUnicodeテキスト描画
 - ✓ OpenTypeと既存のフォントの違い
 - ✓ ISO/IEC 14496-22が規定する範囲
- OFFとISO/IEC 9541との整合性
 - OFFの書体管理情報
 - ✓ OFF-9541間で相互変換すると異なる書体になる例
 - ✓ ISO/IEC 9541-1, -2の更新
 - OFFのグリフ記述方式
 - ✓ ISO/IEC 9541-3更新の必要性
 - ✓ TrueTypeフォントの特徴
 - ✓ TrueTypeヒント情報の処理系依存性
 - ISO/IEC 9541-4の作成
- 更新作業の状況
- 今後の課題
 - 2008年2月の進展

ISO/IEC 9541規格群

- フォント資源の汎用的な情報交換規格
 - Part 1: フォント資源の体系と属性
 - Part 2: フォント資源情報の交換様式
 - Part 3: フォント資源中のグリフ形状表現様式
- フォント資源の情報交換用の規格
 - ファイルのデータ構造は規定しない
 - 文字符号とは独立
 - ✓ 印字処理に必要な属性は規定する
 - ✓ 符号化文字列の印字処理は規定しない

ISO/IEC 9541のスコープ

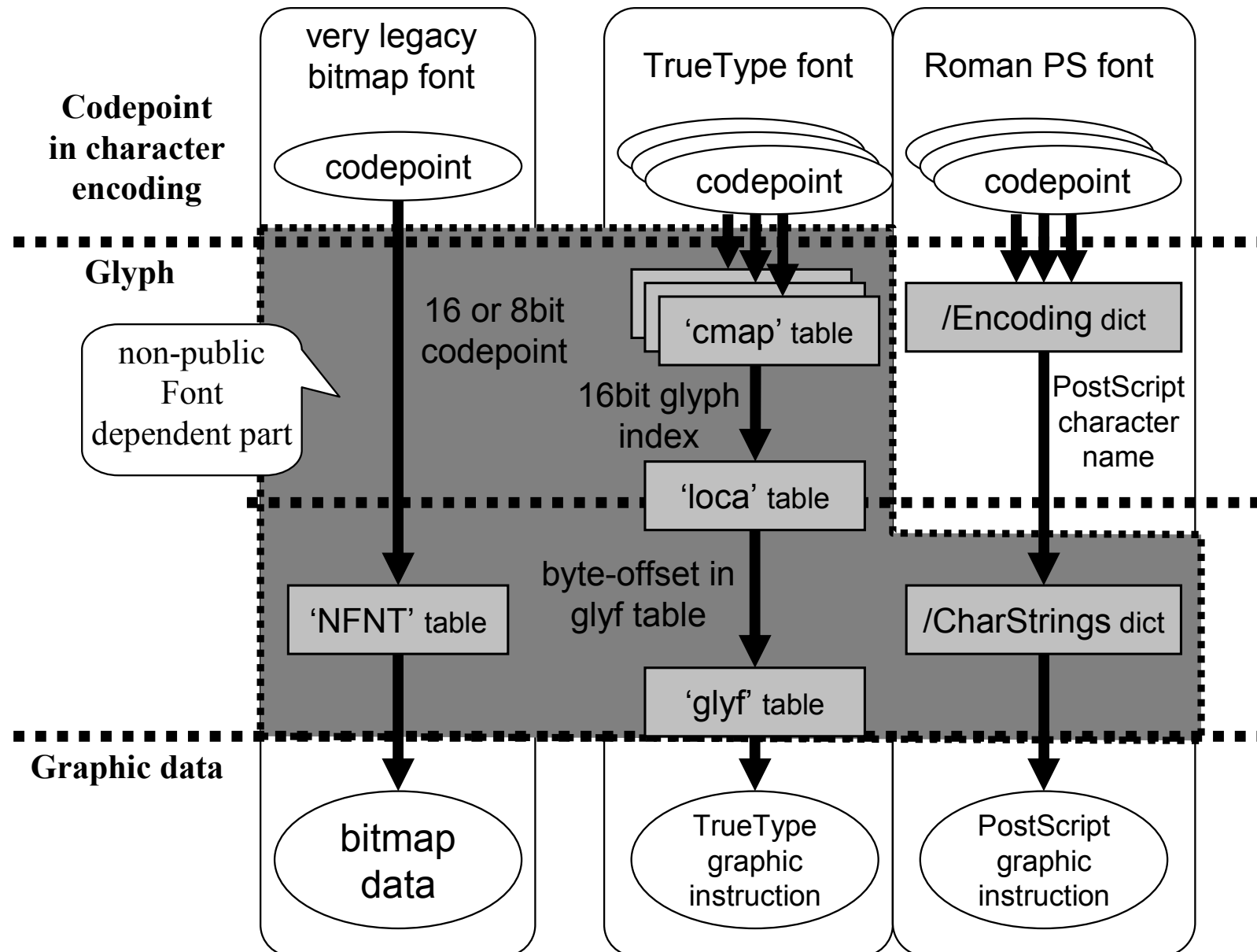


- 出版文書・事務文書はネットワークを通じて交換される
 - フォントを指示できるがフォント資源を含まない
- フォントファイル自体は交換しない
 - フォント処理系が異なる可能性も考慮
- フォントファイルが持つ情報のうち、処理可能な部分がフォント資源となる
 - フォントファイルのデータ構造は規定しない
- フォント資源が管理すべき情報を標準化し、交換可能とする


フォントのデータ構造を規定する規格


- 画像データの一部で、フォント単体の規格ではない
 - ISO/IEC 10180 (Standard Page Description Language)
 - ✓ SPDL処理系から見たフォント資源のデータ構造を規定
 - ファイル形式ではない(他の処理系とは交換しない)
 - Adobe PostScript Type1フォント互換
 - ✓ 符号化文字列の印字処理について、PostScript Level 1と同等の機能を定義
 - 日中台韓の16ビット文字符号には対応可能
 - グリフ集合は定義しない
 - ISO/IEC 16500
 - ✓ DAVIC規格で採用されたBitstream社のPortable Font Resource (PFR)を採用
 - ✓ 8ビットまたは16ビット文字符号に対応可能
- 最も簡単な書字モデルに基づく
 - 1個の文字符号に対し1個の表示図形(グリフ)が一意的に決定される
 - 複雑なテキスト描画が必要な言語の場合、符号化文字列だけでなく、グラフィクス機能の支援が必要
 - ✓ PostScript, PDF, PCLなどの最終出力形式と同様

従来の欧文用のフォントの構造

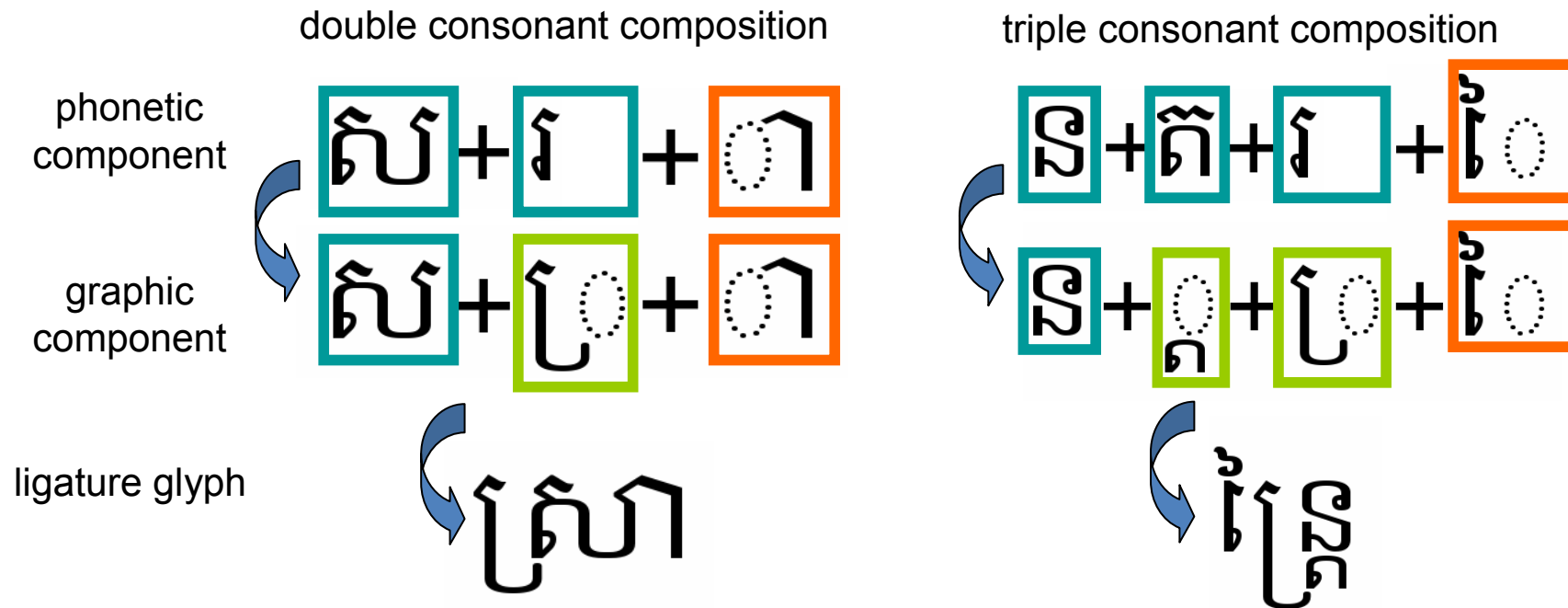


複雑な書字モデルの例 (クメール文字)

 consonant
(base script)

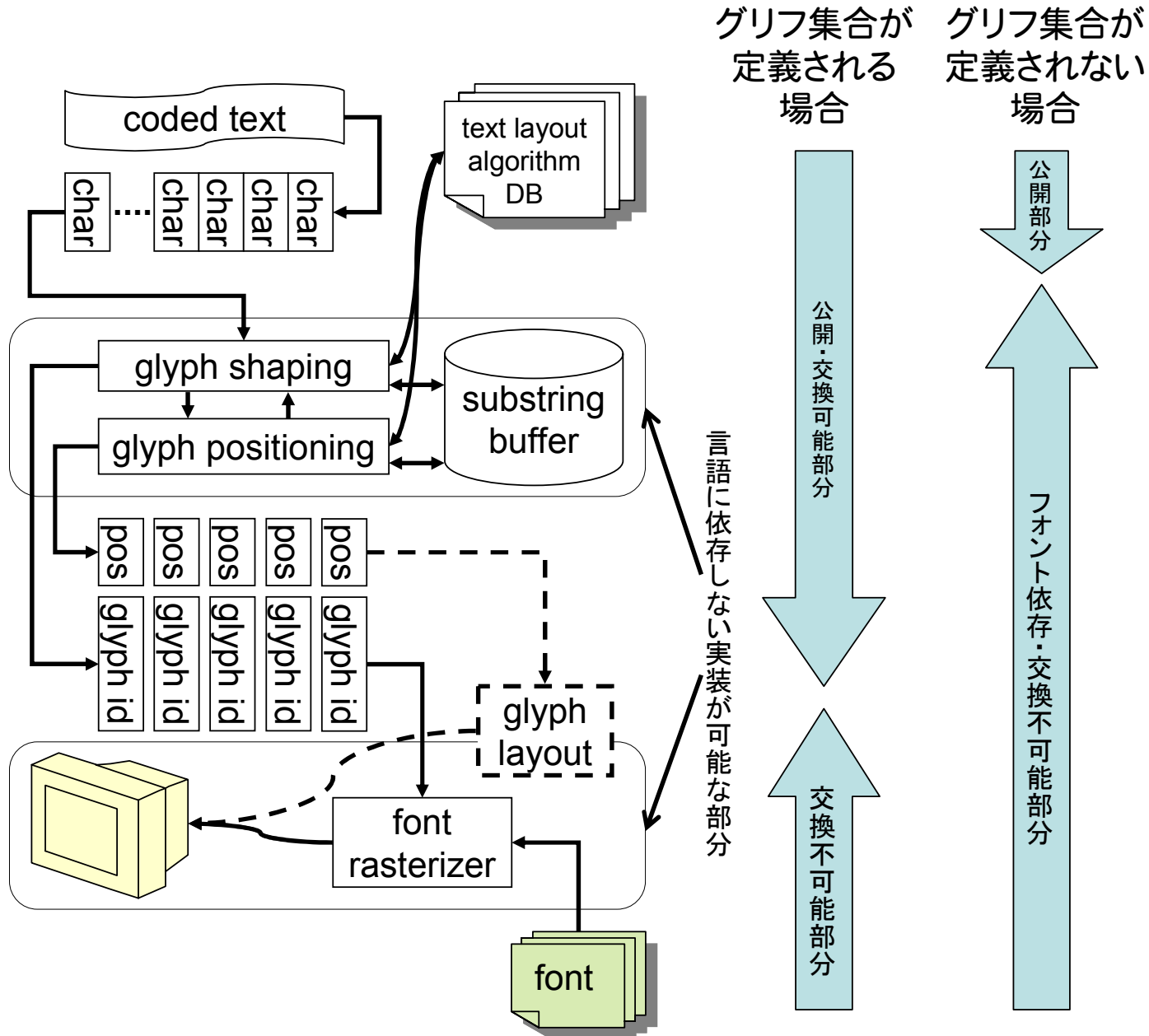
 sub-consonant
(sub script)

 vowel sign



- 符号化の順序と表示の順序は同一でない
- 文字符号からグリフが一意的に定まらない
(周囲の文字符号列により変化する)
- 重ね打ちで合字を作ることが困難

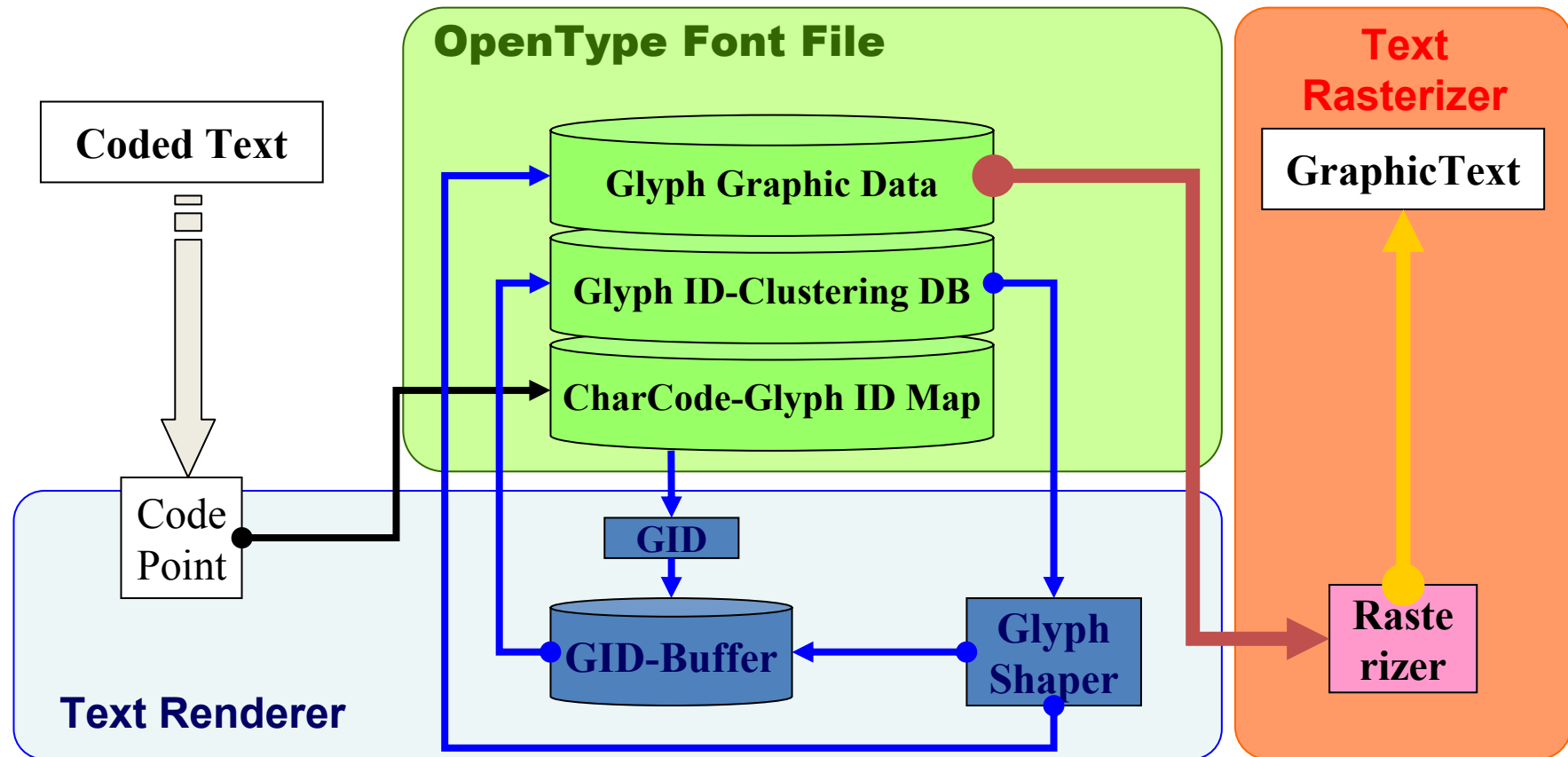
Unicodeテキスト描画処理の流れ



SC29/WG11によるOpenType標準化

- 動画像データへのフォント埋め込み手法は既に標準的なもの(PFR)が存在する(PFR)が、国際化機能が不十分である
 - 書字モデルが単純なため、Unicodeテキストとフォント資源だけで表示できるのはごく一部の言語のみ(ローマ字や漢字など)
 - MPEG-4では動的にテキストを生成したり、ユーザからの入力を表示する必要があり、あらかじめ画像化する手法では制限が多い
- Unicodeテキストの描画資源として広く用いられているOpenTypeをMPEG-4の標準フォント資源とする
 - OpenType 1.4規格を技術的な変更なしにISO/IEC 14496-22 Open Font Formatとして標準化
 - ✓ OpenTypeがデファクト標準であることを考慮したISO化なので、非互換なデジュール標準を定めるべきではないとの判断
- ISO/IEC 16500と同様にフォント資源を動画像データへ埋め込む目的であり、抽象的なフォント資源指定による情報交換は保証していない
 - ISO/IEC 14496-22はOpenType 1.4と同一にするため、他のMPEG-4規格を参照しておらず、スコープが不明確

OpenTypeを用いたUnicodeテキスト描画



- グリフのラスタイザの他に、テキストレンダラを持つ
- グリフの置換・レイアウト規則はOpenTypeフォントが持つが、どの規則を用いるかはテキストレンダラが決定する
 - OpenTypeフォントがテキストレンダラに指示する方法はない
 - テキストレンダラの動作はOpenTypeフォントの規格外

OpenTypeと既存のフォントの違い

- グリフ列に対するパターンマッチによるグリフ置換ができる
 - 標準的な符号化方式で合成文字を表示できる
(外字を用いない)
- 特定のグリフ列パターンに対するグリフ配置情報を格納できる
 - 様々な書字モデルでのグリフ配置・空間調整に対応できる
(従来のフォントは一次元的なペアカーニング)

ISO/IEC 14496-22が規定する範囲

- フォントのデータ構造

- フォント全体の管理用情報
- 文字符号-グリフ番号対応情報
- 各グリフのメトリクス
- グリフ列の置換情報および二次元配置情報
- グリフ形状のビットマップ表現

- グリフ形状のベクトル表現については私的規格を参照している

- Microsoft TrueType instruction part 1 & 2
- Adobe Compact Font Format Specification
- Adobe Type2 CharStrings Specification

- フォント処理系の動作は規定しない

OFFとISO/IEC 9541との整合性

- OFFをISO/IEC 9541体系で用いる場合、
 - OFFデータからISO/IEC 9541準拠の属性をどう定義すればいいのか
(フォント指定ができない)
 - ✓ 同一の管理情報を各処理系用に個別に格納している場合がある
(OS/2用, PCL用, PostScript用など)
 - ✓ 名称が同じでも定義がISO/IEC 9541と異なるものがある
(Italicなど)
 - OFFデータのグリフ形状表現方式がISO/IEC 9541-3では未定義
(グリフ形状の交換ができない)
- OFFをOpenType 1.4同等とした背景を考慮し、
ISO/IEC 9541側で整合性をとる

OFFの書体管理情報

- 形式が規定されない管理情報
 - フォント作成者、バージョン、ライセンス
- 抽象化しない数値
 - 最大幅、最小幅
 - グリフ毎のメトリク
 - 斜体を合成する際の傾き、下線を付加する場合の距離
 - アウトラインデータの解像度
- 特定の出力装置のための補助パラメータ
 - PostScript用
 - PCL用
- フォント資源の抽象的指定のために数値化されたパラメータ
 - 書字方向
 - IBM書体分類
 - Panose書体分類
 - フォントに含まれる文字集合
 - ✓ ISO/IEC 10646部分集合による分類
 - ✓ 現地語化コードページ番号による分類

OFF-9541間で相互変換すると異なる書体になる例

MORT Spabefg ITC Korinna IBM::Clarendon::Monotone	<ul style="list-style-type: none"> ● large x-height ● monotone stems 	MORT Spabefgo ATF STYMIE ISO::Serifs::Square::Monotone
MORT Spabefg ITC Lubalin IBM::Slab::Monotone	<ul style="list-style-type: none"> ● large x-height ● a square transition between strokes/serifs ● serifs/strokes of equal weight 	<ul style="list-style-type: none"> ● serifs that are approximately the same weight (thickness) as the stems ● no fillets or brackets connecting the stems and serifs
MORT spabefg Linotype Candida IBM::Slab::Humanist	<ul style="list-style-type: none"> ● large x-height ● a square transition between strokes/serifs ● serifs of lighter weight than stroke 	
MORT spabefgo Linotype Rockwell IBM::Slab::Geometric	<ul style="list-style-type: none"> ● large x-height ● a square transition between strokes/serifs ● serif/strokes of equal weight ● geometric design 	
MORT spabefg Linotype Serifa IBM::Slab::Swiss	<ul style="list-style-type: none"> ● large x-height ● a square transition between strokes/serifs ● serif/strokes of equal weight ● emphasized white space 	
MORT Spabef IBM Courier IBM::Slab::Typewriter	<ul style="list-style-type: none"> ● large x-height ● a square transition between strokes/serifs ● serif/strokes of equal/moderate thickness ● geometric design 	

ISO/IEC 9541-1, -2の更新

- IBM書体分類、Panose書体分類ともISO/IEC 9541-1 Amd. 1の分類と互換性がない
 - 少なくともPanose書体分類値はグリフの画像解析により得られるものなので、ISO/IEC 9541-1 Amd. 1の書体分類とのマッチングは取れない
 - OFFの書体分類値を透過的に扱うための属性情報をISO/IEC 9541側で追加する
 - IBM書体分類値はISO/IEC 9541を参照としているが、例示書体に不適切なものがあり相互変換ができない
 - 対応づけられない書体について、ISO/IEC 14496-22側でISO/IEC 9541の参照を削除することを要請

OFFのグリフ記述方式

●ビットマップデータ

- ISO/IEC 14496-22内で定義されている
(Appleの非公開の圧縮方式は除外)

●アウトラインデータ

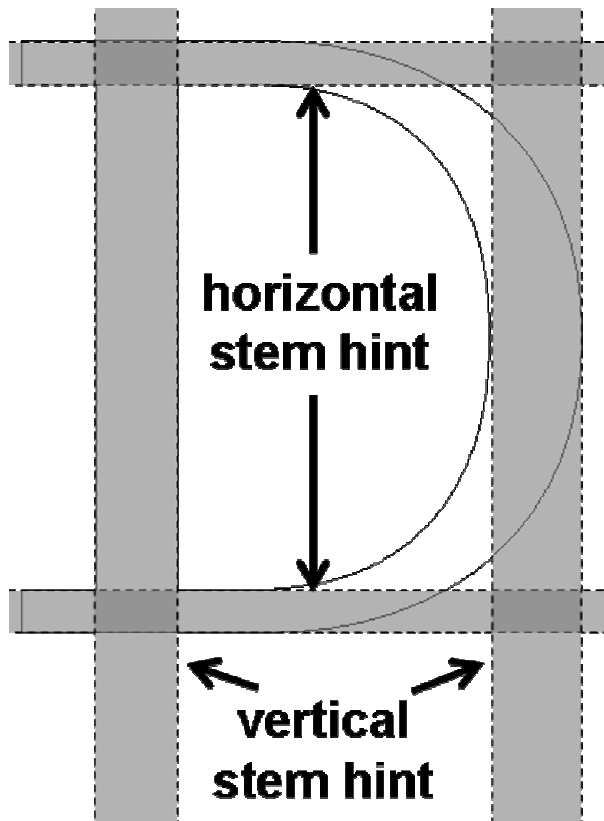
- TrueType描画命令群
 - ✓ MicrosoftのTrueType instruction仕様を参照
(Apple仕様とは若干異なり、最近も拡張されている)
- Adobe Type2 CharString命令群
 - ✓ Adobe Technical Note #5177仕様を参照
(ISO/IEC 9541-3のType1描画命令群の拡張)

ISO/IEC 9541-3更新の必要性

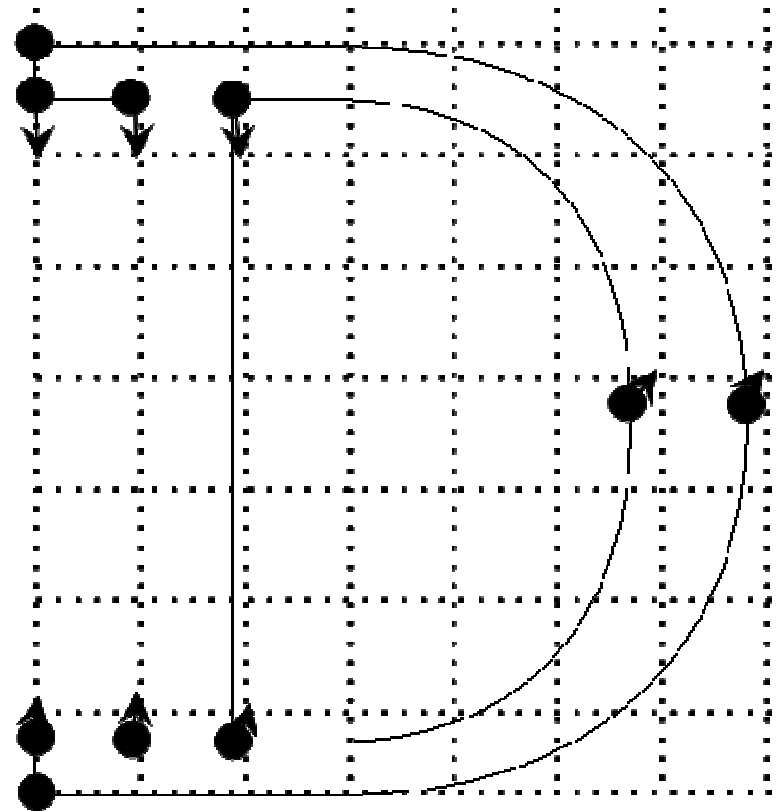
- OFFの2つのアウトライン記述方式は、他のISO標準でも利用されているが、標準化されていない
 - PDF/X, PDF/Aなどの埋め込みフォント
- TrueType instruction setの仕様は固定していないので、標準化が必要
 - Apple仕様とMicrosoft仕様が既に異なる
 - Microsoft仕様は現在も拡張されている
- Adobe Type1 CharString描画命令群が既に標準化されており、Type2 CharString描画命令群への追従が必要
 - 命令群のシンタックスは異なるが処理系はほぼ同一

TrueTypeフォントの特徴

PS Type1 hinting



TrueType hinting



- PostScriptフォントに出力デバイスのピクセルを考慮したヒント情報を与えることが難しい(ヒント情報の反映手法は実装依存)
- TrueTypeフォントではスプライン曲線の制御点全てにグリッド点への量子化ヒントを与えることができる(低解像度デバイスでアウトライン表示が可能)
 - 制御点の量子化までは規格で定義される(PSフォントとの相互変換は困難)

TrueTypeヒント情報の処理系依存性

	IPA明朝		平成明朝W3		平成ゴシックW3	
処理系A	7pt	漢字	7pt	漢字	7pt	漢字
	8pt	漢字	8pt	漢字	8pt	漢字
	9pt	漢字	9pt	漢字	9pt	漢字
	10pt	漢字	10pt	漢字	10pt	漢字
	10.5pt	漢字	10.5pt	漢字	10.5pt	漢字
	11pt	漢字	11pt	漢字	11pt	漢字
	12pt	漢字	12pt	漢字	12pt	漢字
処理系B	7pt	漢字..	7pt	漢字..	7pt	漢字..
	8pt	漢字..	8pt	漢字..	8pt	漢字..
	9pt	漢字+	9pt	漢字+	9pt	漢字+
	10pt	漢字+	10pt	漢字+	10pt	漢字+
	10.5pt	漢字+	10.5pt	漢字+	10.5pt	漢字+
	11pt	漢字+	11pt	漢字+	11pt	漢字+
	12pt	漢字	12pt	漢字	12pt	漢字

ISO/IEC 9541-4の作成

- ISO/IEC 9541の数値型属性と、OFFの書体管理用の数値の対応が未定義
 - たとえばiso-standard-9541-minsizeとOFF headテーブルにおけるxMin, yMinは同じものを指すが単位系は異なる
- OFFは同一の数値情報を異なる表現形式で複数の場所に格納している
 - たとえばiso-standard-9541-weightに対応する数値をOFFは5箇所に格納しているので、どの数値を用いてISO/IEC 9541の属性を定義すべきか不明確

→ 対応関係や優先順位の定義が必要

更新作業の状況

●昨年末のISO/IEC JTC1/SC34総会で 下記の投票を行うことが決定

- ISO/IEC 9541-1, -2の更新 → FPDAM
- ISO/IEC 9541-3の更新
 - ✓ Adobe Type2 CharString描画命令群 → PDAM
- ISO/IEC 9541-4の作成 → CD

●3/14に投票締め切り

今後の課題

● OFFの構造の未標準化部分

- CFFテーブルは外部の私的規格を参照
 - ✓ Adobe Type1 CharString描画命令を用いて良いのか？
- ISO/IEC 9541はファイル構造を規定するものではない

● グリフ命名規則の標準化

- 置換後のグリフに対する命名規則が標準化されていない
 - ✓ 従来のTrueTypeフォントではローマ字についてはAdobe Glyph List、その他についてはISO/IEC 10646コードポイント値が使われるが、置換後のグリフ名については標準的なものがない
 - ✓ ISO/IEC 9541ではグリフ名での指定ができるが、OFFのグリフにはどのように命名されているか予測できない

● 復号処理の標準化

- ISO/IEC 14496-22はOFFデータ構造の標準化であって、処理系の動作は標準化していない
 - ✓ ISO/IEC 14496-22だけではUnicodeテキストの描画は保証できない

2008年2月の進展

- MicrosoftがOpenType 1.5仕様の確定作業中
 - ISO/IEC 10646最新版とUnicode 5.0への準拠
 - Unicode異体字タグのサポート
 - SC29/WG11メンバーがISO/IEC 14496-22:2008として平行作業中
 - SC34のリエゾンコメントがSC29/WG11経由で転送され、IBM書体分類におけるISO/IEC 9541の間違った参照がOpenType仕様から削除された
- ISO/IEC 14496-4:Amd. 26としてOFFファイルおよび処理系の実装水準の規定が作業中
 - 水準の区分にいくつかの問題が見つかり、SC34/WG2からリエゾンコメントを送付
 - ✓ PDFに埋め込まれたフォント資源を適切に表現できない
 - ✓ ビットマップフォントのみの最も単純な水準を適切に表現できない