

Multi-Organ Segmentation with Missing Organs in Abdominal CT Images

Miyuki Suzuki¹, Marius George Linguraru², and Kazunori Okada¹

¹ Department of Computer Science, San Francisco State University

² Sheikh Zayed Institute for Pediatric Surgical Innovation
Children's National Medical Center, Washington DC
{miyukis,kazokada}@sfsu.edu, mlingura@cnmc.org

Abstract. Currently, multi-organ segmentation (MOS) in abdominal CT can fail to handle clinical patient population with missing organs due to surgical resection. In order to enable the state-of-the-art MOS for these clinically important cases, we propose 1) automatic missing organ detection (MOD) by testing abnormality of post-surgical organ motion and organ-specific intensity homogeneity, and 2) atlas-based MOS of 10 abdominal organs that handles missing organs automatically. The proposed methods are validated with 44 abdominal CT scans including 9 diseased cases with surgical organ resections, resulting in 93.3% accuracy for MOD and improved overall segmentation accuracy by the proposed MOS method when tested on difficult diseased cases.

1 Introduction

Multi-organ segmentation (MOS) has recently become popular toward improving overall segmentation accuracy when segmenting a set of organs located nearby, enabling comprehensive computer-aided diagnosis (CAD) of various multi-focal abdominal diseases [1–10]. In this paper, we investigate how such MOS can be extended to a patient population with missing organs due to surgical resections. Without considering this population, MOS cannot be applied to a number of important clinical applications such as follow-up studies of surgical treatment and cancer recurrence in abdomen. Despite this clinical importance, however, current MOS solutions are not designed to handle such cases with irregular anatomy. A common process in various MOS methods is to fit an atlas of normal organ anatomy to an image to be analyzed. When analyzing a case with missing organs, regardless of atlas formats (i.e., static [3], probabilistic [2, 4, 5, 8, 9], or geometric [4, 6–8, 10]), MOS can fail to segment other intact organs because of 1) mis-match of the atlas' part corresponding to the missing organs to nearby non-targets and 2) post-surgical organ shifts. Fig.1(a) illustrates such a failure case with a missing right kidney where the liver (red) shifted downward into the cavity caused by the removed kidney and a part of the liver was incorrectly identified as kidney (cyan).

Addressing the above issue, this paper presents two novel contributions to improve the current atlas-guided MOS solutions. First, we propose an automatic

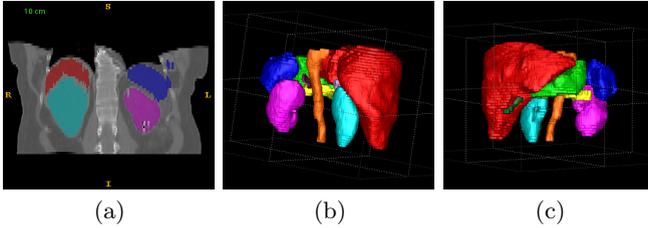


Fig. 1. Illustrative examples of a) segmentation failures (part of the liver is incorrectly labeled as kidney) and b,c) ten modeled organs. Red: liver, blue: spleen, cyan: r-kidney, magenta: l-kidney, yellow: pancreas, orange: aorta, dark green: gall bladder, purple: l-adrenal, lavender: r-adrenal, green: stomach.

missing organ detection (MOD) solution based on testing abnormality of data-driven features computed from the 4D spatio-intensity Gaussian mixture model (GMM) fitted to data. Three probabilistic features, capturing post-surgical organ motions, organ-specific intensity homogeneity, and their linear combinations, are proposed and compared. Such automatic MOD allows us to handle clinical scan data more robustly even when previous medical history information is missing or corrupted in patient record or DICOM tag [11]. Second, we present an atlas-guided MOS solution for 10 abdominal organs that automatically handles missing organs by incorporating the MOD solution to an atlas-guided maximum-a-posteriori (MAP) algorithm proposed in [4]. These proposed methods are validated with 44 abdominal CT scans, including 9 diseased cases with two common surgical resection procedures of splenectomy (spleen removal) and nephrectomy (kidney removal). Our experimental results demonstrate advantages of the proposed MOS method such that a correct MOD improves overall segmentation accuracy on average when dealing with the difficult diseased cases. The issue of handling missing organs in abdominal MOS is scarcely addressed in the literature. To the best of our knowledge, there is no previous studies that proposed an abdominal multi-organ segmentation with automatic missing organ handling.

2 Method

2.1 Atlas-Guided MAP Multi-Organ Segmentation

An atlas-guided MOS method proposed by Shimizu et al. [4] is adopted in this study as our base MOS method. This method employs the MAP estimation of organ label $l \in \{1, \dots, L\}$ over 4D spatio-intensity feature vector $\mathbf{v} = (x, y, z, I(x, y, z))$: $\hat{l} = \operatorname{argmax}_l p(\mathbf{v}|l)p(l)$. The prior $p(l)$ is modeled by a standard probabilistic atlas [2, 9]. The atlas $A_l(\mathbf{x}) \in [0, 1]$, $\mathbf{x} = (x, y, z)$, is built by registering K training images of normal anatomy to a fixed reference image I_R with a size-preserving affine registration then computing a probability map for each of L modeled organs by counting manually segmented organs. The likelihood $p(\mathbf{v}|l)$ is modeled by an extended GMM $p(\mathbf{v}) = \sum_{l=1}^L \sum_{n=1}^N \alpha_l(n) \mathcal{N}(\mathbf{v}; \mathbf{u}_l, \Sigma_l)$ where N

denotes the number of voxels and the mixing weights $\alpha_l(n)$ are defined over each voxel n . To segment organs in a new image, the image is first registered to I_R using affine transformation followed by B-spline non-rigid registration [12]. From the K training images, a normal spatio-intensity model $(\mathbf{u}_{\mathbf{v}l}, \mathbf{\Sigma}_{\mathbf{v}l})$ for each organ l is also computed where $\mathbf{u}_{\mathbf{v}l}$ and $\mathbf{\Sigma}_{\mathbf{v}l}$ are the mean and covariance of feature vectors of the organ l . Initialized by this normal spatio-intensity model, $p(\mathbf{v})$ is fit to the new image using the EM-algorithm [13], yielding the patient-specific likelihood estimate $\{\hat{p}(\mathbf{v}|l)\}$. Additionally, the fitted GMM yields data-driven estimate of organ center and associated covariance $(\hat{\mathbf{u}}_{\mathbf{x}l}, \hat{\mathbf{\Sigma}}_{\mathbf{x}l})$ for each organ l .

2.2 Automatic Missing Organ Detection (MOD)

When fitting the GMM $p(\mathbf{v})$ to an image I^{mo} with missing organs, normal components in $p(\mathbf{v})$ corresponding to missing organs will be fitted to arbitrary non-target structures located nearby. Exploiting this observation, we propose a data-driven MOD by analyzing this EM model fitting error. Three probabilistic measures of missing organs, F_l , G_l , and H_l , are derived by testing abnormality of organ features estimated from the GMM fitting result with respect to respective normal models, as described below.

The first measure F_l indicates the probability of organ l to be missing by quantifying how abnormal the estimated organ center \mathbf{x} is spatially. Geometry of abdominal organs varies due to a) inter-subject variation, b) post-surgical organ shifts, c) postures and d) pathology. To account for the first two factors, the normal spatial models of organ centers are constructed separately for cases with normal anatomy and with different patterns of missing organs due to varying surgical resection procedures. Let MO and NA denote sets of training samples with and without missing organs, respectively. And $MO_{t=1,\dots,T}$, denotes training samples for the t -th surgical organ resection procedure where T indicates the total number of resection procedures considered and $MO = \bigcup_t MO_t$. Then normal anatomy model M^{na} and missing organ model M^{mo} are defined by the following sets of normal distributions,

$$M^{na} = \{M_l^{na}\} = \{\mathcal{N}(\mathbf{x}; \mathbf{u}_l^{na}, \mathbf{\Sigma}_l^{na}) | l = 1, \dots, L\} \quad (1)$$

$$M^{mo} = \{M_{tl}^{mo}\} = \{\mathcal{N}(\mathbf{x}; \mathbf{u}_l^{mo_t}, \mathbf{\Sigma}_l^{mo_t}) | t = 1, \dots, T, l = 1, \dots, L\} \quad (2)$$

where $(\mathbf{u}_l^{na}, \mathbf{\Sigma}_l^{na})$ denote the mean and covariance of the center location for organ l averaged over NA , while $(\mathbf{u}_l^{mo_t}, \mathbf{\Sigma}_l^{mo_t})$ denote those averaged over MO_t for the t -th resection procedure.

We define F_l given M^{na} and M^{mo} as follows,

$$\begin{aligned} F_l &= 1 - p(\mathbf{x}|\theta_l) \\ &= \min(1 - \mathcal{N}(\mathbf{x}; \mathbf{u}_l^{na}, \mathbf{\Sigma}_l^{na}), \{1 - \mathcal{N}(\mathbf{x}; \mathbf{u}_l^{mo_t}, \mathbf{\Sigma}_l^{mo_t})\}_{t=1,\dots,T}) \\ &= 1 - \max(\mathcal{N}(\mathbf{x}; \mathbf{u}_l^{na}, \mathbf{\Sigma}_l^{na}), \{\mathcal{N}(\mathbf{x}; \mathbf{u}_l^{mo_t}, \mathbf{\Sigma}_l^{mo_t})\}_{t=1,\dots,T}) \end{aligned} \quad (3)$$

where $\theta_l = ((\mathbf{u}_l^{na}, \mathbf{\Sigma}_l^{na}), \{(\mathbf{u}_l^{mo_t}, \mathbf{\Sigma}_l^{mo_t})\})$. This measure yields high value when the estimated organ center does not follow trends captured in none of the known normal anatomy or surgical procedure-specific models.

The second measure G_l examines the abnormality in texture pattern homogeneity. For each organ l , a binary mask $B_l(\mathbf{x})$ representing an average shape of the organ is derived from the probabilistic atlas by setting $B_l(\mathbf{x}) = 1, \forall \mathbf{x}$ $A_l(\mathbf{x}) = 1$ and zero otherwise. Using these binary masks, intensity entropy $E_{lm} = -\sum_{i=1}^B p_{lm}(i) \log p_{lm}(i)$ are computed for each organ l in all training samples of NA , where $p_{lm}(i)$ is a B -bin normalized histogram of intensity values sampled under $B_l(\mathbf{x})$ in the m -th sample. For each organ l , the mean and standard deviation of the entropy distribution (E_l^{na}, σ_l^{na}) are computed over $\{E_{lm}\}$, forming a normal model of organ-specific texture homogeneities. To evaluate an organ l , the entropy E_l of the organ is computed by overlaying $B_l(\mathbf{x})$ by aligning its gravity center to the estimated organ center in the new image and sampling intensity values within the mask. Then G_l is defined as an abnormality measure of E_l with respect to the normal model,

$$G_l = 1 - p(E_l | \phi_l) = 1 - \mathcal{N}(E_l; E_l^{na}, \sigma_l^{na}) \quad (4)$$

where $\phi_l = (E_l^{na}, \sigma_l^{na})$.

The third measure H_l is defined as a linear combination of F_l and G_l ,

$$H_l = \beta F_l + (1 - \beta) G_l \quad (5)$$

where $\beta \in [0, 1]$. Finally, missing organs are detected by applying a threshold function to these measures derived for each organ in a new image for arbitrary number of missing organs per case.

2.3 Multi-Organ Segmentation (MOS) with Missing Organs

As a final step, the base MOS method described in Sec 2.1 can be adopted to missing organ cases by discarding the atlas A_l and the spatio-intensity model $\mathcal{N}(\mathbf{x}; \mathbf{u}_{\mathbf{v}l}, \Sigma_{\mathbf{v}l})$ corresponding to missing organs during the model fitting and inference procedures. The entire MOS procedure thus consists of three successive steps: 1) the base MOS, 2) MOD with F_l , G_l , or H_l , and 3) the modified MOS without A_l , $\mathbf{u}_{\mathbf{v}l}$, and $\Sigma_{\mathbf{v}l}$ for the detected missing organs.

3 Experiments

3.1 Data

A total of 44 abdominal CT scans are used in this study. Ten non-contrast thin-slice (1mm) abdominal CT scans of healthy volunteers ($K = 10$) are manually segmented by expert radiologists and used to construct the probability atlas by . The NA set contains 25 contrast-enhanced abdominal CT scans with normal anatomy, while the MO set consists of 9 diseased scans with three types ($t = 1, 2, 3$) of surgical organ removal: i) 5 splenectomy cases (spleen removed), ii) 3 nephrectomy cases (right kidney removed), and iii) 1 splenectomy and

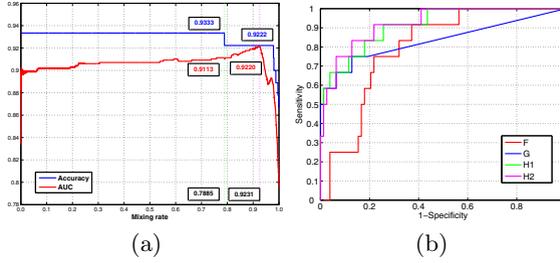


Fig. 2. Quantitative validation of the proposed MOD. (a) Maximum accuracy and AUC values with various mixing rate β for computing the H_l measure. Green and magenta dotted-lines denote β values that yield the maximum accuracy and the maximum AUC, respectively. (b) ROC analysis of MOD with four different measures: red, F_l , blue, G_l , green, H_1 with $\beta = 0.789$, and pink, H_2 with $\beta = 0.923$.

nephrectomy case (spleen and left kidney removed). Each scan consists of $512 \times 512 \times 50$ voxel slices with 5mm slice thickness stored in Mayo analyze format. CT scanners from various manufacturers are used to acquire this dataset with the ISOVUE 300 contrast agent. Ten abdominal organs ($L = 10$) are considered in this study: aorta (AO), gallbladder (GB), left/right adrenal glands (LA,RA), liver (LV), left/right kidney (LK,RK), pancreas (PN), spleen (SP), and stomach (ST). For validation, segmentation ground-truth is generated for 9 NA and 9 MO cases by expert researchers with ITK-Snap tool. Fig.1(b,c) illustrate some examples.

3.2 Results

Leave-one-out cross validation is performed to validate the performance of the proposed MOD method on the MO set. For each of the three measures, we evaluated 50,000 different detection thresholds with a fixed interval between 0 and 1 and derived the receiver operating characteristic (ROC) curves. Maximum accuracies ($TP+TN/TP+TN+FP+FN$) with minimum false positive rate was 0.867 and 0.933 for F_l and G_l , respectively. The number of 80 bins ($B=80$) was used to derive G_l . For H_l , we evaluated 50,000 different mixing rate β values with a fixed interval between 0 and 1. Fig. 2(a) shows the maximum accuracy and the area under the ROC curve (AUC) computed for various β values. The linear combination did not increase the accuracy measure; the maximum accuracy of 0.933 with highest AUC of 0.911 was found at $\beta = 0.789$ (referred as H_1). On the other hands, the overall maximum of AUC with 0.922 was found at $\beta = 0.923$ with slightly decreased accuracy of 0.922 (referred as H_2). Fig. 2(b) shows the ROC curves for F_l , G_l , H_1 , and H_2 , clearly demonstrating the advantage of the proposed linear combination measure. AUC values for F_l and G_l were 0.795 and 0.834, respectively.

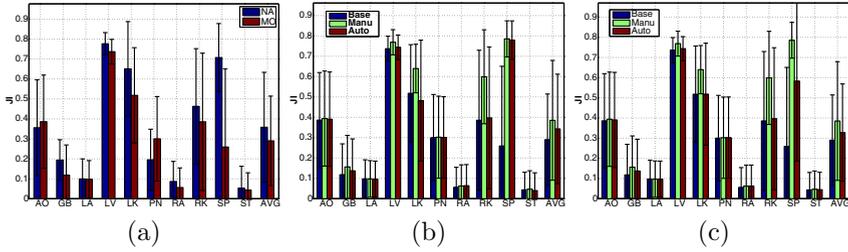


Fig. 3. Average Jaccard index computed for 10 abdominal organs, comparing different MOS methods and datasets. (a) Performance by the base MOS method (*Base*) for normal anatomy (*NA*) and missing organs (*MO*) cases. (b) Comparison of the base and the proposed methods with automatic (*Auto*) and manual (*Manu*) MOD with $\beta = 0.923$ (*H2*) on *MO*. (c) With $\beta = 0.789$ (*H1*).

We next evaluate the proposed MOS method with the missing organ cases. Fig. 3(a) shows organ-wise segmentation accuracy of the base MOS method [4] in Jaccard index (*JI*) on the nine normal anatomy *NA* and the nine diseased *MO* cases as baseline. Liver, left kidney, and spleen have relatively high accuracy. Segmentation of adrenal glands and gall bladder is challenging because they are very small and their shape varies widely. Stomach also yields very low *JI* because its shape and intensity is extremely variant. For most organs, the accuracy for *MO* cases is lower than that for *NA*. The accuracy for spleen and left kidney in *MO* is largely lowered due to missing them in some cases of *MO*. Not only missing organ itself but even neighboring organ, liver, is influenced by right kidney missing such that the bottom of liver is segmented as right kidney that causes the lower accuracy of *MO* liver.

Fig. 3(b) and (c) compare the accuracy in *JI* for the base and the proposed MOS methods with automatic and manual MOD on *MO* cases with the two versions of H_l measures with $\beta = 0.789$ (*H1*) and $\beta = 0.923$ (*H2*), respectively. The manual MOD specifies which organs are missing according to the ground-truth labels. In both versions, the MOS with manual MOD (*Manu*) performed better than the base method (*Base*), demonstrating the proof-of-concept of our approach in improving segmentation accuracy by explicitly considering missing organs. Our proposed fully-automatic method (*Auto*) outperformed *Base* on average for both versions, although accuracy was lowered from that of *Manu* due to the MOD errors. For spleen, the proposed *Auto* method largely improved *Base* in both versions. The accuracy for the left kidney was slightly improved with $\beta = 0.789$, and that for liver and right kidney was also slightly improved by the both versions of *Auto*.

Fig. 4 shows four illustrative examples for segmenting splenectomy cases (missing spleen). In these examples, spleen (blue), as well as other organs such as gallbladder (dark green) and pancreas (yellow), are fully or partially resected surgically. The examples show that the missing organs are correctly detected by our method and existing neighboring organs, such as left kidney (magenta), is also correctly segmented despite its post-surgical organ shifts. Fig. 5 compares

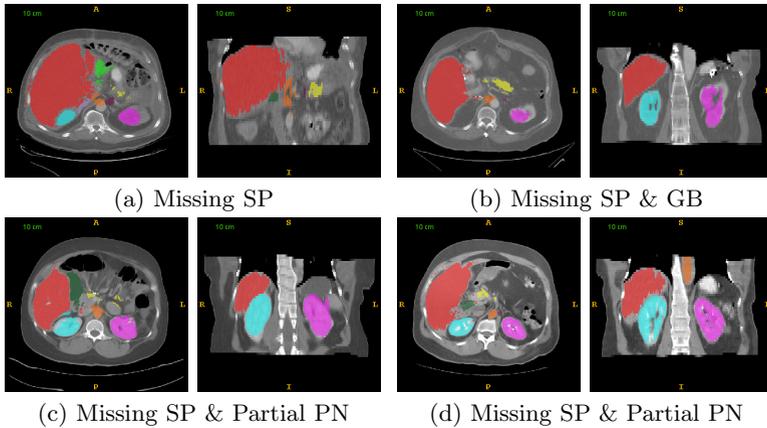


Fig. 4. Four illustrative splenectomy examples of MOS by the proposed *Auto* method. Spleen (blue) is missing in these examples.

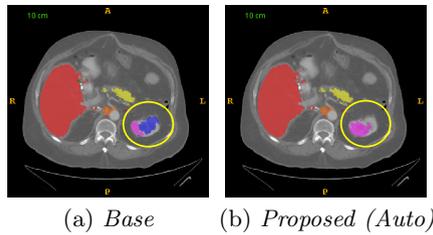


Fig. 5. Segmentation comparison for neighboring organ; (a) the missing spleen (blue) is incorrectly placed inside the left kidney; (b) the improved segmentation

the segmentation results by the base and proposed methods in the splenectomy example in Fig. 4(b). The base method without MOD falsely segments a large part of left kidney (magenta) as (missing) spleen (blue) as shown in Fig. 5(a). Fig. 5(b) clearly shows that the correct MOD of spleen leads to much better segmentation of the neighboring kidney.

4 Conclusions and Discussion

This paper presented novel methods for automatic MOD and atlas-guided MOS that handle missing organs. Our experimental results are promising in that 1) high accuracy of MOD was observed even with the limited number of missing organ cases used in training and 2) the proposed MOS improved the average *JI* accuracy, demonstrating the advantage of our MOD-MOS approach. As our future work, more missing organ cases and surgical resection procedures must be included to further our study in 1) post-surgical organ shifts in finer details and 2) MOD and MOS of partially resected organs that were not addressed in this paper. Finally, we plan to improve the accuracy of our MOS solution, especially

for those difficult organs, by improving our atlas and GMM models, as well as by refining the discontinuous segmentation results by using our results to initialize other graph-based/contour-based segmentation solutions.

References

1. Kobatake, H.: Future CAD in multi-dimensional medical images - project on multi-organ, multi-disease CAD system. *Computerized Medical Imaging and Graphics* 31, 258–266 (2007)
2. Park, H., Bland, P.H., Meyer, C.R.: Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans. Medical Imaging* 22, 483–492 (2003)
3. Zhou, Y., Bai, J.: Multiple abdominal organ segmentation: An atlas-based fuzzy connectedness approach. *IEEE Trans. Info. Tech. in Biomed.* 11, 348–352 (2007)
4. Shimizu, A., Ohno, R., Ikegami, T., Kobatake, H., Nawano, S., Smutek, D.: Segmentation of multiple organs in non-contrast 3D abdominal CT images. *Int. J. CARS* 2, 135–143 (2007)
5. Okada, T., Yokota, K., Hori, M., Nakamoto, M., Nakamura, H., Sato, Y.: Construction of Hierarchical Multi-Organ Statistical Atlases and Their Application to Multi-Organ Segmentation from CT Images. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I. LNCS*, vol. 5241, pp. 502–509. Springer, Heidelberg (2008)
6. Seifert, S., Barbu, A., Zhou, S.K., Liu, D., Feulner, J., Huber, M., Suehling, M., Cavallaro, A., Comaniciu, D.: Hierarchical parsing and semantic navigation of full body CT data. In: *Proc. SPIE Conf. Medical Imaging* (2008)
7. Yao, J., Summers, R.M.: Statistical Location Model for Abdominal Organ Localization. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part II. LNCS*, vol. 5762, pp. 9–17. Springer, Heidelberg (2009)
8. Linguraru, M.G., Pura, J.A., Chowdhury, A.S., Summers, R.M.: Multi-organ Segmentation from Multi-phase Abdominal CT via 4D Graphs Using Enhancement, Shape and Location Optimization. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010, Part III. LNCS*, vol. 6363, pp. 89–96. Springer, Heidelberg (2010)
9. Linguraru, M.G., Sandberg, J.K., Li, Z., Pura, J.A., Summers, R.M.: Atlas-based automated segmentation of spleen and liver using adaptive enhancement estimation. *Medical Physics* 37, 771–783 (2010)
10. Liu, X., Linguraru, M.G., Yao, J., Summers, R.M.: Organ Pose Distribution Model and an MAP Framework for Automated Abdominal Multi-organ Localization. In: Liao, H., Edwards, P.J., Pan, X., Fan, Y., Yang, G.-Z. (eds.) *MIAR 2010. LNCS*, vol. 6326, pp. 393–402. Springer, Heidelberg (2010)
11. Guld, M.O., Kohonen, M., Keysers, D., Schubert, H., Wein, B.B., Bredno, J., Lehmann, T.M.: Quality of DICOM header information for image categorization. In: *Proc. SPIE Conf. Medical Imaging*, vol. 4685, pp. 280–287 (2002)
12. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Trans. Medical Imaging* 18, 712–721 (1999)
13. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy Stats. Soc. Series B* 39, 1–38 (1977)